DOCUMENT RESUME

ED 350 305 TM 017 253

AUTHOR	Baker, Eva; And Others
TITLE	Report on a Validity Study of the Tennessee Career Ladder System.
PUB DATE	May 90
NOTE	240p.
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Tests/Evaluation Instruments (160)
EDRS PRICE	MF01/PC10 Plus Postage.
DESCRIPTORS	*Career Ladders; *Elementary School Teachers; Elementary Secondary Education; Evaluation Methods; Evaluators; Inferences; Measures (Individuals); Program Evaluation; Program Implementation; *Secondary School Teachers; State Programs; *Teacher Evaluation; *Validity
IDENTIFIERS	*Tennessee; *Tennessee Career Ladder Program; Validity Research

ABSTRACT

To fulfill a mandate of the Tennessee State Legislature, the extent to which the Tennessee Career Ladder System (CLS) is valid for making inferences about the professional competence of teachers was investigated. Evidence of the extent to which the content of the evaluation procedures reflects sound theory, empirical findings, and best practices, and statistical indicators of the extent to which procedures contribute to accurate scores and inferences were examined. Chapter 1 of this paper is an introduction. Chapters 2 through 5 provide support for content validity as it pertains to CLS instruments and procedures. Chapter 6 describes statistical indicators supporting the validity of inferences drawn from the CLS. Conclusions are presented in Chapter 7. Responses from 1,990 teachers in the CLS were used to obtain teacher feedback about the CLS observation instrument and procedures. Results suggest that technically sound and professionally recognized procedures were used to identify domains and competencies assessed by the CLS instruments. Instrument content and evaluator selection processes were judged appropriate, as were evaluator training, rater reliability, and program implementation. Overall, the CLS staff appears to have done a stellar job. Ten tables present study data. Seven appendixes provide information about the CLS system, national education goals, the validity studies, and the instruments used. Thirty-seven references are included. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Doris Redfield

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED350305

Report on a Validity Study of the Tennessee Career Ladder System

Submitted May, 1990

Validation Team

Dr. Eva Baker, Team Chair
Director, National Center for the Study of Evaluation, Standards, and Student Testing
UCLA Graduate School of Education

Dr. C. Thomas Kerins
Manager, Student Assessment Section
Illinois State Board of Education

Dr. Robert L. Linn
Co-Director, National Center for the Study of Evaluation, Standards, and Student Testing
University of Colorado

Dr. Doris Redfield
Project Director, Center for the Study of Evaluation
UCLA Graduate School of Education

Dr. Paul D. Sandifer
Director, Office of Research
South Carolina State Department of Education

TM017253

BEST COPY AVAILABLE

TABLE OF CONTENTS

Page

EXECUTIVE SUMMARY vi

TECHNICAL REPORT

- Chapter I: Introduction 1
- Chapter II: Content Validity—Classroom Observation Instrument 12
- Chapter III: Content Validity—Professional Development and Leadership Summary 31
- Chapter IV: Content Validity—Dialogue Instrument 43
- Chapter V: Content Validity—Questionnaires 73
- Chapter VI: Statistical Indicators of Validity 86
- Chapter VII: Conclusions 104

REFERENCES 114

TABLES 1-10 120

APPENDICES

- A: Tennessee Code Annotated, Section 49-5-5103(4) of the Comprehensive Education Reform Act 130
- B: National Educational Goals 133

C: Outline of Supporting Material Available for Proposed Activities to Validate Career Ladder Evaluation System 136

D: Proposal—Tennessee Career Ladder System Validation Study 173

E: Validation Team Organizational Meeting Agenda 181

F: Career Ladder System Instruments 183

G: How Information is Obtained about Competencies and Indicators 224

LIST OF TABLES

Page

Table 1: Weighted Coefficient Alpha Reliability Estimates 120

Table 2: Multitrait-Multimethod Correlation Matrix for Data Subscores in Domains I Through IV 121

Table 3: Correlations of Observation, Dialogue, Student, and Skills Test Subscores with the Combined DATA Score Within Domains I Through IV 122

Table 4: Correlations of Data, Principal, and Consensus Subscores with the Domain Score within Each Domain 123

Table 5: Correlations of Domain Scores with the Composite Scores 124

Table 6: Illustration of Estimating Career Ladder Level Based on Four of Five Domain Scores 125

Table 7: Number of Applicants for Career Ladder Level III, for Levels II or III, and Percent of Applicants Attaining Those Levels by Year and Race 126

Table 8: Number of Applicants for Career Ladder Level III, for Levels II or III, and Percent of Applicants Attaining Those Levels by Year and Gender 127

Table 9: Estimated Percent of Applicants for Career Ladder Levels II or III Who Would Have Attained the Level Applied for if Alternative Composites Were Used 128

Table 10: Career Ladder Application and Attainment for Tennessee Teachers Who Have Received Other Teaching Awards 129

EXECUTIVE SUMMARY

This report is a manifestation of the requirement, in Tennessee Code Annotated, Section 49-5-5103(4) of Tennessee's Comprehensive Educational Reform Act, for evidence of the extent to which the Career Ladder System (CLS) is valid for making inferences about the professional competency of teachers. To fulfill this required activity (Appendix A), we have investigated two kinds of evidence: (a) evidence of the extent to which the content of the evaluation procedures reflect sound theory, empirical findings, and best practice and (b) statistical indicators of the extent to which the procedures contribute to accurate "scores" and, hence, inferences about teacher competency.

Parameters of the Study

The principles which guided our investigation include the following:

1. The entire system, rather than individual instruments or procedures, should be the basis for judging the validity of inferences about professional competency. Sound measurement theory and practice suggest that no single data source can consistently provide an accurate indicator of a complex construct such as instructional competency. The validity of inferences rests with the system of indicators and not with individual measures which make up the system.
2. Our charge has been to investigate evidence for the validity of the CLS; it has not been to evaluate the CLS. Validity

studies are context and purpose specific, focussing on the extent to which the inferences made from particular data sets are appropriate, meaningful, and useful with regard to particular circumstances. Evaluation studies, on the other hand, are policy oriented. While our findings may have implications for policy, our study has not been policy driven.

3. Since validity is context and purpose specific, it has not been possible for us to determine the extent to which the Tennessee CLS may be used to make valid inferences for unintended purposes. The results of this study are meant within a particular context and should not be overgeneralized.
4. The limits of this study should affect the interpretation and generalization of results. These limitations include (a) our exclusive reliance upon extant data, (b) our exclusive focus on teachers, and (c) our study of group differences in percent passing to the exclusion of bias.

Summary of Findings

The overall findings yielded by our investigation are highlighted below.

Development Activities. Existing documentation justifies the conclusion that technically sound and professionally recognized procedures were used to identify the domains and competencies assessed by the CLS instruments. The evidence further suggests that the processes and methods used to modify and refine the various instruments and procedures over time have been appropriate, politically sensitive, and based on the results of data analyses, expert judgment, and feedback from CLS participants.

Content of the Instruments. Given appropriate training and implementation, results yielded by the instruments should contribute to valid inferences regarding teaching quality. The instruments' competencies and indicators are based on effective teaching research, consensus by Tennessee teachers, expert judgment, and the reasoned approval of the State Certification Commission (SCC).

Evaluator Selection. The criteria and procedures for selecting evaluators are basically sound. The application criteria include teaching experience, subject matter and grade level expertise, writing ability, and communication skills. All of these seem both reasonable and necessary basic qualifications which, along with an adequate training program, should yield competent evaluators. The current selection process is reasonable and responsive to the need to fairly interview large numbers of candidates within short and intense timeframes. The process appropriately requires that (a) all applications be read by two Department of Education staff and rated according to the qualifying criteria and (b) highly rated candidates be personally interviewed by two members of the staff.

Evaluator Training. There is strong evidence that the content of training in the Dialogue, Observation, and Professional Development and Leadership Summary (PDLS) instruments is appropriate and adequate. We observed the training to be of sufficient duration and depth to provide a basis for reliable use of the instruments and related procedures. By comparison, training in Questionnaire administration is minimal. The need for extensive training is minimized by the straightforward instructions for Questionnaire administration contained in the Evaluator Manual (e.g., Tennessee State Department of Education, 1988-89). Nonetheless, actual administration allows for discretion on the part of the evaluator--discretion which seems reasonable but which may also influence the reliability of results.

Rater Reliability and Implementation. To be certified as evaluators, trainees must meet the criteria for reliability. The criteria are based on 80% agreement with pre-established expert judgment. Hence, it may be concluded that certified evaluators are able to reliably apply the Dialogue, Observation, and PDLS instruments. The extent of interrater reliability based on actual classroom observations, rather than on training tapes, however, is not known. Also unknown is the extent to which a candidate's writing ability is undoubtedly and unavoidably confounded with the quality of the activities assessed by the PDLS. Several controls have been built into the system to help compensate for such confounding (e.g., independent readings and ratings by two evaluators who must reach consensus).

The Student and Principal Questionnaires are reliable as demonstrated by their internal consistency coefficients. However, the validity of the Principal Questionnaire should be interpreted with caution since the principal's ratings are individual, identifiable from other data sources, and ultimately known to the candidate.

Group Differences in Passing Rates. The data indicate that black teachers have been somewhat over-represented at Career Ladder Level I and under-represented at Levels II and III. This finding, regardless of possible reasons such as missing data, held for every year analyzed (1986-89). There are also gender differences in passing rates. Female applicants, particularly for Level III, have a higher success rate than do their male counterparts. Alternate composite scores, based on systematic exclusion of individual domain scores, were used to estimate failure rates by race and sex. The alternate composites did not substantially change the group differences in passing rates.

There is relatively little variation in passing rates across the major geographic regions of the state. However, relatively large differences in the percentage of teachers at Career Ladder Levels II and III exist with respect to per capita income.

External Criteria. Teachers judged as outstanding on grounds other than the CLS are more likely to apply for CLS candidacy than their unrecognized peers. They are also more likely to achieve

Level II or Level III status, indicating that the CLS is a valid indicator of teaching excellence.

Psychometric Properties. Insofar as applicable, internal consistency reliability has been established for each of the CLS domains: Planning, Teaching Strategies, Evaluation, Classroom Management, and Professional Development and Leadership. The magnitudes of the five domain coefficients are reasonable, especially since high internal consistency is not a requirement for effective measures based on composites of relatively heterogeneous subscores.

Multitrait-multimethod methodology provided weak evidence of convergent validity among domain scores and no evidence of differential validity. These findings are not notably unusual and suggest the importance of using more than one evaluation method to arrive at composite scores as well as the importance of basing decisions regarding CLS attainment levels on composite, rather than individual domain, scores.

The high correlations between scores within each domain and their corresponding domain scores suggest that the Principal Questionnaire and consensus scores are technically unnecessary. However, their inclusion in the system adds to the credibility of the process and does not detract from its validity.

The domain scores for all domains except Professional Development and Leadership have relatively high intercorrelations. These findings suggest that relative standings across domains should be interpreted cautiously.

Analyses of composite scores that were based on the systematic elimination of single domain scores and weighted combinations of the remaining scores indicates that similar decisions regarding attainment levels would result. Hence, while all domain scores positively contribute to the composite score, eliminating single domain scores and redistributing the weight of the missing domain score among the remaining domains does not differentially affect attainment decisions about Career Ladder Levels.

Overall Conclusion

Within the confines of the legislation, the enormity of the task, and the available resources, CLS staff have done a stellar job of conceptualizing, developing, and implementing a model teacher evaluation system that allows for valid inferences about teaching quality. Suggestions for further consideration or study are provided at the end of Chapter VII.

Chapter I: Introduction

The appropriate, useful, and meaningful evaluation of education is a complex enterprise. When educational evaluation involves judgments about the competency of personnel, it is additionally and particularly controversial. Nonetheless, public and political pressure to engage in such evaluation has increased throughout the 1980s and, as evidenced by the National Goals (Appendix B) emanating from the President's 1989 Educational Summit with the Governors, the need for valid educational assessments will remain with us into the 21st century.

Tennessee has been a pioneer in developing ways to meet the needs for educational personnel evaluation. In January, 1983, Governor Lamar Alexander proposed a ten-point "Better Schools Program." His proposal seemed particularly timely, given the April, 1983 release of A Nation at Risk (National Commission on Excellence in Education, 1983). A prominent part of Governor Alexander's proposed program was the development of a Career Ladder evaluation system for teachers and administrators. Given the paucity of appropriate methods or models for meeting the need-at-hand, this developmental approach was defensible.

For the most part, career ladder systems assume that the quality of schools impacts the quality of learning, that school quality is largely dependent upon the quality of the personnel responsible for delivering instruction, and that incentives can motivate instructional improvements. Clearly, Tennessee has not been alone in considering the potential of incentive programs for school improvement (Cornett, 1990).

In the Spring of 1983, an interim certification commission was appointed to develop Tennessee's Career Ladder System (CLS). With the passage of the Comprehensive Education Reform Act (CERA) of 1984, and the accompanying legislative funding, the CLS became a reality.

During August and September, 1984, the Interim Certification Commission (ICC) and the State Board of Education approved implementation of the CLS for teachers. As documented elsewhere (Appendix C), intensive effort was put into the development of the System prior to its initial implementation in 1984-85. Several thousand Tennessee teachers participated in the preparation of various parts of the System; several hundred more assisted with field testing efforts.

In keeping with the law, i.e., Tennessee Code Annotated, Section 49-5-5103(4), one of the CLS-related duties of the State Certification Commission (SCC) is as follows: " . . . report to the State Board of Education and the Legislative Oversight Committee annually on the validation and testing of evaluation criteria, . . . to assure that educators receive a fair, unbiased, and objective determination of professional competency." A copy of Tennessee Code Annotated, Section 49-5-5103(4) is provided in Appendix A.

Purpose

This report is a manifestation of the CERA requirement for evidence of the extent to which the CLS is valid for making inferences about the professional competency of teachers. To meet this requirement, two kinds of evidence are necessary: (a) evidence of the extent to which the content of the evaluation procedures reflect sound theory, empirical findings, and best practice and (b) statistical indicators of the extent to which the procedures contribute to accurate "scores" and, hence, inferences about teacher competency. In preparing this report, we investigated both kinds of evidence.

Guiding Principles

The following principles guided our investigation:

Principle #1

The entire system, rather than individual instruments or procedures, should be the basis for judging the validity of inferences about professional competency. Evaluation systems incorporate numerous measures in order to more accurately and consistently assess complex outcomes such as instructional competency. For example, instructional competency may be dependent upon a teacher's general knowledge, knowledge in a specific subject matter area, knowledge of how to effectively teach particular subject matter, attitudes toward teaching and learners, etc. Sound measurement theory and practice suggest that no single data source can consistently provide an accurate indicator of a complex construct such as instructional competency. Validity thus rests with the inferences made on the basis of the system; it does not rest with the individual instruments which make up the system.

Principle #2

Our charge has been to investigate evidence for the validity of the CLS; it has not been to evaluate the CLS. Validity studies are context specific and focus on the extent to which the inferences made from particular data sets are appropriate, meaningful, and useful. Hence, in this study, our focus has been on the extent to which scores yielded by the CLS accurately differentiate between "better" (Level II) and "best" (Level III) teachers. Evaluation studies are more policy oriented. While our findings may have implications for policy, our study has not been policy driven. In the spirit of this principle, we have chosen to look at the feasibility of the System without judging the desirability of the philosophies and processes underlying its legislation.

Principle #3

Since validity is context and purpose specific (AERA, APA, NCME, 1985), it is not possible for us to determine the extent to which the Tennessee CLS can be used to make valid inferences for unintended purposes. While a number of different inferences may be based on particular data sets and while there are many ways to gather evidence in support of particular inferences, statements about validity must always be restricted to particular inferences. Statements of validity, then, refer to specific uses of assessment data. Therefore, the results of this study are meant within a particular context and should not be overgeneralized.

Principle #4

It is important to recognize the limitations of this study insofar as they may affect the interpretation and generalization of results. Limitations include the following:

- We have relied on extant data only.
- We have focused on teachers only.
- We have addressed the issues of differential passing rates for majority and minority group candidates but we have not directly addressed the extent to which the system may be biased. Bias concerns the fairness of using assessments for certain minority groups and should not be equated with differential passing rates.

For purposes of this study, and as illustrated by the study proposal in Appendix D, operating within these limits was agreed to be reasonable and justifiable. If follow-up study is warranted, and if resources are available, some of these limits may be eliminated or minimized.

Issues and Questions Addressed by the Study

With the above principles in mind, the issues addressed by our study included the following:

- the relationship between the CLS and its supporting theoretical, empirical, and practical bases. Aspects of the CLS under consideration included the instruments, training procedures, implementation procedures, scoring, and the assignment of standards-based attainment levels.
- the over-time relationship between teachers' participation in the CLS and certain teacher characteristics (e.g., race, sex).
- the extent to which scoring and weighting procedures accurately reflect levels of teaching excellence.
- the relationship between CLS data and other available indicators of teaching excellence (e.g., outstanding teacher awards).

The more specific questions addressed by our study are built around the above issues and are provided as they apply to the information contained in the remaining chapters of this report. They are also outlined in Appendix D. The following section of this chapter describes the processes we used to conduct the validation study. These processes were guided by the previously described principles and were designed to address the issues listed above.

Process

The Validation Team met first in Nashville on January 24 and 25, 1989. All, or a majority of the Team, have since met in March 1989, June 1989, and February, 1990. The Team consists of Dr. Eva Baker, University of California at Los Angeles (UCLA); Dr. C. Thomas Kerins, Illinois State Board of Education; Dr. Robert Linn, University of Colorado; Dr. Doris Redfield, UCLA; and Dr. Paul Sandifer, South Carolina State Department of Education. Dr. Baker is Team Chair. Team members' resumes are on file with the Tennessee State Department of Education.

The purpose of the January 1989 meeting was for the Team to consult with Tennessee Career Ladder staff in order to define the ensuing validation study and to identify potential data sources. Additionally, the Team agreed on workplan parameters, including a preliminary timeline and task assignments. The resultant workplan was submitted, in proposal form, to the Tennessee State Department of Education in March, 1989. The overall purpose and structure of the study has remained constant, as proposed; however, to accommodate the Team's desire to remain sensitive to on-going needs and the developmental nature of the CLS, the fine-grain, detailed aspects of the proposal have been been treated formatively. Hence, the proposed workplan, with the knowledge and approval of CLS officials, has functioned as a working document throughout the course of the validation study. Copies of the January meeting agenda and the most recent draft of the proposed workplan appear in Appendices D and E.

Briefly, as a result of the January meeting, the Team divided tasks into two general categories: those associated with empirical evidence of content validity and those associated with statistical evidence of validity and fairness. Drs. Kerins, Redfield, and Sandifer assumed primary responsibility for the former; Drs. Baker and Linn for the latter. At the January meeting, the Team, along with CLS staff, agreed upon the following procedural principles:

- The first year of the validation study should focus on examining extant data rather than on collecting new data.
- While individual Team members would assume primary responsibility for particular tasks, the Final Report would be a product of consensus among all Team members.
- A statement of the purpose of the study would appear in the frontpiece of the Final Report, as agreed upon by Tennessee State Department of Education officials and the Team.
- In examining the extent to which the CLS contributes to valid inferences about the differentiation between Level II and III teachers, the Final Report should address issues of content validity, equity and access, and criterion-related validity. The issues to be explored include the following: (1) content validity evidence of the relationships among CLS Levels, the CLS instruments and procedures, and the theoretical/empirical bases

underlying the various instruments and procedures; (2) statistical evidence of the longitudinal relationships among teachers' participation in the CLS, various teaching contexts, and certain teacher characteristics (e.g., race, sex); (3) statistical evidence of the longitudinal relationships among teacher characteristics, teaching contexts, and CLS Levels; (4) the content validity and statistical evidence of the extent to which scoring and weighting procedures accurately reflect levels of teaching excellence; and (5) the relationship between CLS data and other available indicators of teaching excellence.

Since the January, 1989 meeting in Nashville, team members have worked independently to carry out their assigned tasks while maintaining regular communication via computer, fax, correspondence, and telephone. For the Subteam working on content validity issues, primary tasks have included extensive review of extant documents, data, and files as well as direct observation of evaluator training sessions in the Observation, Dialogue, and Professional Development & Leadership Summary instruments. For the Subteam working on statistical indicators of validity, primary tasks have included examination and secondary analyses of extant data for purposes of addressing questions of criterion validity and fairness.

Documents reviewed by the team included Evaluator Training Manuals (e.g., Tennessee State Department of Education, 1989-89), Career Ladder Teacher Orientation Manuals (Tennessee State Department of Education, 1984-90), minutes of State Certification Commission meetings, Sounding Board summaries, and various commissioned or staff-developed background and technical papers. Appendix C provides a complete list of pertinent documents which are available in the Tennessee Department of Education files.

In March 1989, while attending the Annual Meeting of the American Educational Research Association (AERA), the Team met as a whole and in task-related subgroups. CLS staff members were available as needed in order to clarify issues emerging as a result of the Team's preliminary work.

In June 1989, those members of the Team attending the Annual Meeting of the Education Commission of the States (ECS) in Boulder, met to clarify and discuss data files and analyses required for ensuing statistical analyses. Again, CLS staff familiar with the extant data were available for consultation as needed.

In February 1990 the Team met, in Nashville, with CLS staff and State Department officials. The intent of the meeting was for the Team to discuss their preliminary findings and obtain clarification and feedback for purposes of finalizing their analyses, conclusions, and recommendations.

Throughout the process, Tennessee CLS staff have been most cooperative in helping the Team understand the chronology of events and in locating extant data. However, they have not been directly or otherwise inappropriately involved in the validity study.

Summary

This report is in response to the validation activity required by Tennessee Code Annotated, Section 49-5-5103(4) of Tennessee's Comprehensive Educational Reform Act (Appendix C). It examines evidence of the extent to which the CLS may be used to make valid inferences about teaching competency (i.e., CLS levels II and III). Chapters II-V provide evidence of content validity as it pertains to CLS instruments and procedures. Chapter VI describes statistical indicators supporting the validity of inferences emanating from the System. Conclusions based upon the available evidence and our examination of it are provided in Chapter VII. To allow each chapter to be understood on a stand-alone basis, there is some redundancy across chapters.

Chapter II: Content Validity—Observation Instrument

The Classroom Observation Instrument is but one of several data sources that are used in forming judgments concerning whether the Career Ladder System (CLS) participants should be classified as "better" (level II), or "best" (level III) teachers. Since inferences concerning the quality of teaching are not made solely on the observation data, the issue of the validity of inferences is moot. There is, however, a need to examine the content of the observation instrument to determine whether the behaviors that are observed can reasonably be expected to assist in differentiating between "better" and "best" teachers. Additionally, it is necessary to examine the manner in which the instrument is applied to determine whether the data obtained from classroom observations are sufficiently reliable to warrant their use in decision making.

Therefore, the purpose of this section of the report is to review the development and implementation of the Classroom Observation Instrument and the related evaluator training as those components of the CLS relate, as applicable, to each of the following four questions.

1. What evidence exists that, as a whole, instrument development procedures reflected the stated purposes of the CLS and reasonable practice?
2. What evidence exists that the content of the observation instrument reflects specified needs, reasonable practice, and empirical findings?

3. To what extent do training procedures ensure that the instruments are reliably used as intended?
4. To what extent are the instruments and related procedures implemented as intended?

Although the discussion in this section of the report is guided by the four questions, no attempt has been made to treat each of the questions independently since to do so would ignore the interrelatedness of some of the questions and would result in considerable redundancy.

The Observation Process and Instrumentation

The observation process includes three major components: the pre-observation conference; classroom observation using the Observation Instrument; and a post-observation conference. A brief description of each of these is given here to provide a context for the subsequent discussion on instrument development.

The pre-observation conference, according to the *Evaluator Training Manual* (e.g., Tennessee State Department of Education, 1988-89), is the time during which "the observer develops a perspective of the classroom, learners, and lessons which he/she is planning to observe." The Manual further states that the conference provides an opportunity for the observer to establish rapport with the teacher and to identify any factors or conditions which might affect the teacher's performance.

The Observation Instrument (Appendix F) consists of three sections and provides a means for structured observation of teacher behavior. Section One provides a record of instructional interactions, student questions/comments, interruptions, involvement of volunteers, and teacher control.

Section Two includes questions on the number of students in the class; whether the class was working as a total group, as sub-groups, or as individuals; whether any students expressed any lack of understanding and, if so, how the teacher responded; whether the teacher presented correct information; safety or sanitation violations; and whether reasonable sanctions for inappropriate behavior were provided. Again, according to the _Evaluator Training Manual_ (e.g., Tennessee State Department of Education, 1988-89), the lesson being observed is divided into six-minute periods during which the observer makes a record of behaviors in Section One for four minutes and then answers additional questions in Section two for two minutes. This process results in at least seven observation periods during the course of a lesson.

Section Three of the Observation Instrument requires that the observer respond to a set of questions about the quality of observed behavior patterns in categories such as planning and management; use of resources; use of facilities; provision of feedback; appropriateness of communication; and classroom climate. Section Three is completed at the end of the class period.

The current observation procedures require two observations by each of three different observers, for a total of six observations. Since each observation contains as many as seven six-minute periods, there are approximately forty-two observation segments and six sets of questions about the quality of observed behavior patterns that are used in determining a teacher's observation score.

The post observation conference is, according to the _Evaluator Training Manual_ (e.g., Tennessee State Department of Education, 1988-89), to be held as soon as possible after the observations. The conference is intended as a means for the observer to accomplish the following three things:

1. Share insights and perceptions gained during the observations and solicit the evaluatee's comments on these;
2. Share specific, competency related strengths and needs observed;
3. Develop with the evaluatee recommendations for improvement.

Additionally, the conference is used for gathering information concerning the major objectives of the lesson, the methods of presentation, and other information which the observer feels may be important to the observation.

Instrument Development

Based on records maintained in the files of the Tennessee Department of Education, the development of the Observation Instrument has apparently been a continuous process of refinement since the initial "Suggested Criteria for Master Teacher Selection and Evaluation" were developed by Rosenholtz and Smylie and submitted to the Interim Certification Commission (ICC) on April 4, 1983.

The April 4, 1983 version was revised by Rosenholtz and Smylie and resubmitted to the ICC in June, 1983. According to the authors, the revision of the initial criteria was undertaken with two goals in mind: "1) To most accurately incorporate the most consistent findings of educational research that relate teaching practices and behaviors to student learning, and 2) to facilitate the development of a system to accurately, fairly, and objectively evaluate teacher performance along those measures that have been found to most consistently relate to student learning" (Rosenholtz & Smylie, 1983a).

The suggested criteria identified by Rosenholtz and Smylie were, with some minor rewording, reformatted, labeled as "Competencies and Indicators", and submitted to teachers for review and comment. Teachers were given the opportunity to judge each indicator as "appropriate" or "inappropriate" and to offer comments and suggestions for change. Although teachers offered many comments and suggestions, their reactions to the appropriateness or inappropriateness of the indicators was overwhelmingly positive (French, 1983).

Prior to the teacher survey, the Competencies and Indicators had not been classified as to possible sources of data. Consequently, the document which teachers reviewed contained a number of Competencies and Indicators which, although they may ultimately have become a part of some other instrument, were not amenable to observation. This is in no way intended as a criticism of the development process but is intended to point out that several of the instruments/procedures used in the CLS had a common beginning in the "Suggested Criteria" developed by Rosenholtz and Smylie.

Based on the results of the teacher survey and consultation with a number of authorities in measurement and evaluation, including Jason Millman, W. James Popham, Robert Soar, and Jane Stallings, staff of the Department of Education revised the Competencies and Indicators and reduced them in number. In December, 1983, the ICC approved a plan for the development and field testing of assessment instruments for the "Career Teacher Evaluation System" (Interim Commission, 1983).

The approved field test plan included a classroom observation component which was to involve three steps. First, a pre-observation conference involving the teacher and evaluator was to be conducted for the purposes of gathering information necessary to the observer's understanding of what would be happening in the classroom, and to identify any special conditions that might influence the teacher's performance. The second step called for multiple observations by the evaluation team. The plan stipulated that any single observation was to be not less than one hour in duration or, in the case of secondary teachers, not less than one complete class period.

The third, and final, step in the observation process required a post-observation conference between the teacher and the observer following each observation. The stated purposes of the conferences were to give the candidate (teacher) an opportunity to record his/her reactions to the data and to provide a means whereby the evaluator could make recommendations to enable the candidate to improve in areas of need.

Except for the initial identification of competencies and indicators, the original CLS was developed and field tested during the period December, 1983 through July, 1984. Evaluator training was conducted during the summer and fall of 1984 and the System was implemented during the 1984-85 school year.

As a part of their continuing efforts to collect information about the CLS, Department of Education staff sent questionnaires to approximately 3,200 teachers who were candidates for Career Levels II and III during 1984-85. The questionnaire, which related to all aspects of the CLS, included five questions specific to the classroom observation and four questions that were to be answered in regard to each of the three evaluators.

Responses from 1,990 (62%) of the candidates, indicated that the teachers believed that the Observation Instrument and procedures: 1) are adequately described in materials provided to candidates; 2) are fair and objective; 3) provide an effective indicator of teacher competency; and 4) require an appropriate amount of candidate preparation. Results further indicated teachers' belief that preparation for the classroom observations is likely to be beneficial to a teacher's classroom performance. A large majority of the respondents additionally gave the evaluators very favorable ratings and judged them to be effective communicators; fair and impartial; helpful, concerned and understanding; and competent (Career Ladder Retreat, 1985).

In addition to the information obtained through the survey of Career Ladder candidates, the staff obtained reactions to the first year of operating the CLS from "Sounding Boards" of administrators and teachers and through debriefings of the evaluators. In view of the generally positive response to the classroom observation component of the CLS, it is not surprising that only three substantive changes related to it were recommended for implementation in 1985-86. Those changes, as approved by the ICC, were: 1) increasing the number of observations from three to six (two announced and four unannounced); 2) including oral communication as an indicator under "teaching strategies"; and 3) expanding the Career Ladder Teacher Orientation Manuals (e.g., Tennessee State Department of Education, 1984-1990) to include all instrumentation. The ICC also approved some other changes related to format and organization of the instrument (Career Ladder Retreat, 1985).

Minutes of the August 19, 1986 meeting of the State Certification Commission (SCC) provide information concerning approved changes in the CLS for the 1986-87 school year. As in prior years, proposed changes in the System were based on feedback from CLS candidates and Sounding Boards, and an analysis of the prior year's data.

Although several changes are described for various components of the System, the minutes indicate that the Observation Instrument and procedures were to remain unchanged except for the addition of items to be field tested (State Certification Commission, 1986). However, the "Teacher Communication" section of the Instrument used in 1985-86 contained four Indicators, i.e. "speaks clearly", "uses vocabulary appropriate to audience", "organizes information", and "uses grammar correctly", that were not included in the 1986-87 version of the instrument.

Minutes of the June 19, 1987 meeting of the SCC include approved CLS changes for 1987-88. According to the minutes, the only changes in the Observation Instrument were the deletion of Vocational Education items which related to student organizations and placement, addition of items to be field tested, and the incorporation of items which were field tested in 1986-87 (State Certification Commission, 1987). A comparison of the Instruments for 1986-87 and 1987-88 confirmed that those were the only changes made. Similarly, a comparison of the Instruments used in 1987-88 and 1988-89 revealed only minor changes which were apparently related to the field testing of new items.

Although the refinement of the Instrument has been a continuous process, the basic content of the Observation Instrument has remained very stable since it was first used in 1984-85. This stability in content is probably attributable, in large measure, to the fact that, in the initial stages of development, only those Competencies and Indicators that 70% or more of the teachers rated favorably were retained (State Certification Commission, 1987).

Two of the four questions posed at the beginning of this section as a basis for examining the Observation Instrument relate to its content and the procedures used in developing it. Based on information obtained through a review of documentation made available by the Tennessee Department of Education, it appears that the development procedures were reasonable and, under the time constraints, quite thorough. Teachers, those primarily affected by the System, were given an opportunity to react to the initial list of proposed competencies and indicators and, as indicated earlier, only those items to which at least 70% of the teachers reacted positively were retained. Subsequent revisions/ refinements of the Instruments including the Observation Instrument were based, in part, on comments/suggestions from teachers who were candidates for Career Levels II or III. The content of the Instrument seems consistent with the research on teacher effectiveness and, not surprisingly, is generally consistent with teacher evaluation systems that have been developed in the past decade. Although verbiage may vary to some extent across systems, most focus, as does the CLS, on planning, delivering, and evaluating instruction, with communication skills and classroom management seen as prerequisite to effective instruction. Although Career Ladder decisions are not based solely on the classroom observations, the content of the Observation Instrument is such that, with appropriate training of observers and appropriate implementation, the results of the observations should contribute toward valid inferences about teacher quality.

Evaluator Qualifications and Selection

Whether the Observation Instrument is utilized as intended is largely dependent on two factors: 1) The basic qualifications of the observers/evaluators; and 2) the adequacy of the training program and follow-up/monitoring procedures. We turn now to these two factors and begin by examining the qualifications of the observers/evaluators by reviewing the selection process and criteria.

The criteria and process for selecting evaluators seem to have evolved in much the same manner as the Observation Instrument, i.e., begin with what appears to be a basically sound approach and refine it on the basis of experience and changing needs. The initial criteria for selecting teacher evaluators were approved by the ICC during its meeting of May 1-2, 1984. The criteria included, but were not limited to, full-time employment as an evaluator; a minimum of twelve years of teaching experience; a reputation among peers as an outstanding teacher; demonstration of emotional stability; ability to work under pressure; appropriate grade level and content expertise; and successful completion of the training program. General criteria which were applicable to the entire pool of evaluators included appropriate representation by race and sex; representation by region of the state; and a provision that at least 80% of the evaluators selected be practicing teachers but that the applicant pool could include, among others, staff of the Department of Education and University faculty or staff (Interim Certification Commission, 1984a).

The process used for selecting the evaluators from among the pool of qualified applicants is also detailed in the ICC minutes of May 1-2, 1984. Basically, the process required that the Department of Education staff review applications for conformity to initial criteria; randomly select, from among those qualified, a number of candidates equal to twice the number of evaluators needed; interview all randomly-selected candidates by teams consisting of two ICC members and a Department staff member (Interim Certification Commission, 1984a).

This admirable, and somewhat ambitious, plan for interviewing finalists led to a situation in which the Commission was faced with the task of interviewing 210 applicants in a period of five to six days. The ICC minutes of July 20, 1984 reflect approval of a modified selection process which made the staff responsible for reducing the candidate pool to 150-160; reducing the interview team to two members (one Commissioner and one staff person); and restricting the interview to 30 minutes (Interim Certification Commission, 1984b).

For 1985-86, the pool of potential evaluators included the 1984-85 evaluators who met performance standards established by the ICC and those teachers who achieved Career Level III status during 1984-85. The selection process for 1985-86 required that:

1) all applications be read and rated by at least two staff members of the Department of Education; 2) the ratings consider, among other factors, years of teaching experience, writing ability, willingness to travel, recommendations of peers and subordinates, subject area and grade level expertise, and communication skills; and 3) that highly rated candidates be personally interviewed by two members of the staff. This selection process reflected a realistic, and understandable, departure from the practice which had required that the interview teams include a member of the Commission (Interim Certification Commission, 1985).

For the 1986-87 school year, the criteria and the selection process for teacher evaluators were, with one exception, apparently unchanged. The one exception noted was a decrease from twelve to eight in the minimum number of years of teaching experience required (Interim Certification Commission, 1986).

In planning for the 1987-88 year, the SCC and Department staff attempted to reduce program costs by hiring part-time evaluators, assigning them close to home, and regionalizing training. Except for the planned shift toward a significant number of part-time evaluators, the criteria for selection were the same as in the previous year. However, attempts to implement the plan to hire part-time evaluators encountered two significant problems which apparently caused the plan to be abandoned. First, district superintendents were opposed to the plan due to the potential disruption of instructional programs. Second, the Department received only 47 applications for 140 positions. As a result, the SCC continued the practice of employing evaluators for either year-long or cycle-long assignments and continued centralized training in Nashville (State Certification Commission, 1982).

Do the program managers recruit and employ individuals who have the basic qualifications necessary to become trained as evaluators? Apparently. The selection criteria which are emphasized include teaching experience, subject matter and grade level expertise, writing ability, and communication skills. All of these appear to be not only reasonable but necessary basic qualifications which should, in conjunction with an adequate and properly implemented training program, yield competent evaluators.

Observer Training

The observer training program for the CLS is lengthy and comprehensive, involving approximately four weeks of initial training plus additional training during the course of the year for all instruments. Initial training for the classroom observations is approximately 30 hours. The training, which is based on a comprehensive manual, includes lecture, discussion, practice through the use of video tapes, feedback, and qualifying/scoring at the conclusion of the session. The initial "classroom" segment of the training is followed by practice in "live" situations and then additional classroom sessions.

Reliability checks are made on the evaluators during the training process and all trainees must pass the reliability tests before being certified to conduct observations. The reliability checks, which require approximately 70% agreement with "expert" judgment, are made independently on seven sections of the Observation Instrument. To become certified as evaluators, the trainees must receive at least two "reliable" ratings on each section of the instrument.

Although evaluators may be reliable in viewing video tapes in training sessions, this does not assure that they are reliable when conducting observations in a classroom. As a means of monitoring this aspect of the program during the initial years of the CLS, reliability checks were also made on evaluators during their actual work with teachers. Teachers who were being observed, however, assumed that the presence of a second observer indicated that the first observer was unreliable. Due to the negative image resulting from teachers' misunderstanding of the reason for the presence of the second observer, the on-site reliability checks were discontinued. However, according to the Evaluator Training Manual (e.g., Tennessee State Department of Education, 1988-89), reliability is periodically monitored by requiring the evaluators to review and evaluate video tapes.

Although the Evaluator Training Manual provided to each evaluator is comprehensive in its coverage of the instruments and procedures, it is not a "do-it-yourself" program. The actual training, as viewed by this writer, employs many examples, illustrations, tapes, and supplementary materials that are not part of the Manual. Consequently, proper operation of the training program requires individuals with an in-depth knowledge of the CLS, a thorough understanding of the concepts and premises upon which the instruments are based, and expertise in analyzing teacher behavior. In short, the quality of the training is highly dependent upon the quality of the trainer(s). Consequently, the recent loss of experienced and qualified staff who were apparently well-versed in all aspects of the CLS may warrant close scrutiny of the training program in the near future to ensure that it is properly implemented and that the evaluators are properly trained.

As stated at the beginning of this section on the Observation Instrument and the related training, two of the four questions of interest are the extent to which the training procedures ensure that the instrument is reliably used as intended and the extent to which the instrument and related procedures are implemented as intended. Based on a review of training materials, observation of the training program, discussion with trainers, and review of extensive files in the Department of Education, there is strong evidence that the content of the training is appropriate and adequate and that the training is of sufficient duration and depth to provide the basis for reliable use of the Observation Instrument.

Whether the Observation Instrument and related procedures are implemented as intended cannot be answered unequivocally in this report. The only way to ensure that such is the case is to observe the observers. Although one cannot state with certainty that the instrument and procedures have, to this time, been implemented as intended, in the absence of any evidence to the contrary, it seems reasonable to conclude that such is the case.

Areas for Future Consideration

The procedures for using the Observation Instrument require that a post-observation conference be held for the purpose of providing the teacher with immediate feedback on his/her strengths and weaknesses. Essentially, this requires that the observer conduct an "on-the-spot" evaluation of what was observed. This practice is generally recommended in the literature on personnel evaluation and is, no doubt, appropriate when there is a single evaluator involved. However, in the case of the CLS and any other system that depends on multiple observations made by different observers, the practice seems inconsistent with the basic concepts of the system, i.e. multiple data sources, multiple observations, and multiple observers to arrive at an evaluation. Additionally, the practice has the potential for providing the teacher with inconsistent feedback from the different observers. If this happens, it does not necessarily mean that the observers are unreliable since there is the distinct possibility that on different days, observing different lessons, they observed different behaviors. Since the evaluation of a teacher is intended to be based on all the data obtained from multiple sources, the Department of Education and the SCC may wish to reconsider the requirement for a post-observation conference.

A second area for consideration relates to documentation of the decision rules concerning the reliability requirements for observers. Staff who were involved in the development of the CLS understand the decision rules, the manner in which they were determined, and the rationale for them. Some, but not all, of the rules are embedded in computer programs but there appears to be no documentation concerning the development of, or rationale for, the reliability criteria. Even if the original staff were still with the project, the documentation would be desirable to make this aspect of the system public. Since the original staff are no longer with the Department, it is essential that the documentation be developed before the "institutional memory" is lost.

A third, and final, area for consideration is the possible reinstatement of the on-site reliability checks on observers. Although there is no evidence that the observer instrument and procedures are not being reliably implemented, data from on-site checks could provide stronger support for this position.

Summary

The content of the Observation Instrument seems to be consistent with the research on teacher effectiveness and the procedures associated with its use appear to be appropriate and adequate for the collection of reliable information. The observers/evaluators appear to have the basic qualifications necessary to become trained as evaluators. The selection criteria which are emphasized include teaching experience, subject matter and grade level expertise, writing ability, and communication skills. All of these appear to be reasonable and necessary qualifications.

The observer training program is intense and comprehensive. Trainees are provided with a detailed manual for use in the training and the staff who were observed conducting the training did an outstanding job. There is, however, a need for the development of a training package which does not rely so heavily on staff who are no longer associated with the CLS.

In sum, the Observation Instrument, the qualifications of the observers, and the quality of the training should lead to results which can contribute towards making valid inferences about teacher quality.

Chapter III: Content Validity—Professional Development and Leadership Summary

The Professional Development and Leadership Summary (PDLS) is one of several instruments used in the Tennessee Career Ladder's process for making decisions about individual teacher quality. Other instruments include the Dialogue, Observation, and a series of Questionnaires (Elementary Student, Secondary Student, and Principal Questionnaires). The total Career Ladder System (CLS) score, upon which CLS advancement decisions are made, is based on a weighted sum of scores on these instruments plus a Professional Skills Test.

The purpose of this chapter is to examine the content of the PDLS in order to determine the extent to which it can be expected to contribute to accurate Career Ladder decisions. That is, this chapter addresses the extent to which the PDLS contributes to accurate decisions about teacher candidate advancement to CLS Levels II ("better") and III ("best"). The questions guiding this chapter are the same as those underlying the preceding chapter on the content validity of the Observation Instrument:

- What evidence exists that instrument development procedures reflected the stated purposes of the Career Ladder System and reasonable practice?
- What evidence exists that the instrument's content reflects specified needs, reasonable practice, and empirical findings?

- To what extent do training procedures ensure that the instruments can be reliably used as intended?
- To what extent are the instruments and related procedures implemented as intended?

Instrument Development and Content

Purpose and Chronology of Development/Implementation

A review of documents on file with The Tennessee Department of Education (Appendix C) indicates that the PDLS was first used during the second year of the CLS in 1985-86. The purpose of the PDLS is to provide information for estimating the extent to which CLS candidates establish and maintain professional leadership roles.

During 1984-85, the first year of CLS implementation, professional development and leadership Competencies were assessed using information yielded by an Applicant or Candidate Interview, Peer Questionnaires, a Superordinate Questionnaire that was usually completed by the candidate's principal, and a Portfolio or file of teacher-developed information. The intent of the Portfolio was to provide evidence of teacher competency in each of the CLS's five competency areas, including professional development and leadership (Competency V).

In 1985-86, the Portfolio, which had become an unpopular "paperwork nightmare" (Furtwengler, 1987, p. 67), was replaced with the PDLS and a Dialogue Instrument. The Dialogue, which consists of a series of three structured interviews designed to assess teacher planning, teaching strategies, and evaluation Competencies, also served to replace the Candidate Interview which had proved to be "a time-consuming ordeal for the evaluators and a stressful and exhaustive experience for the teacher" (Furtwengler, 1987, p. 68). Decisions regarding the unwieldiness and inefficiency of the Portfolio and Candidate Interview seem justifiable based upon the documentation contained in the Department's files (e.g., Sounding Board data, publications).

Prior to the 1986-87 evaluations, the Certification Commission voted to delete the Peer Questionnaire which had focused on the candidate's leadership in the school and profession. Analysis of Peer Questionnaire data provided by the first two years of the CLS indicated that the Instrument did not differentiate among teachers and that administration of it required the interruption of instruction in some schools.

On these grounds, the decision to eliminate the Peer Questionnaire seems justifiable, especially since the competencies assessed by the Peer Questionnaire are also assessed by the PDLS and the weight of the PDLS in the total scoring scheme was increased accordingly. Hence, in year three of CLS implementation, professional development and leadership were assessed using the PDLS and items 13-15 of the Principal Questionnaire; although the wording and format of the PDLS has been clarified and refined over the years, the content of the PDLS has remained unchanged from its initial implementation in 1985-86. Since 1987-88 the PDLS has been due on the first or "A" visit of the evaluation cycle rather than on the second or "B" visit as previously.

Competencies and Indicators

The identification and content validation of Competencies for CLS inclusion was based on effective teaching research, consensus by practicing teachers, and expert opinion. The Competencies and Indicators were approved by an Ad Hoc Interim Certification Commission, convened in April, 1983 and charged with developing a master teacher evaluation system. Staff to the Commission included an executive director, Dr. Russell French from the University of Tennessee at Knoxville, and personnel from the State Department of Education's Division of Research and Development.

Vanderbilt University was contracted to provide research information on teacher effectiveness and indicators of teacher competencies which affect student achievement. This research formed the bases for the content of the various CLS evaluation instruments. Reports of the Vanderbilt work and the accompanying recommendations appear in the CLS files as listed in Appendix C (e.g., Rosenholtz & Smylie, 1983b).

Additionally, experts were consulted with regard to issues, criteria, and processes associated with the evaluation of teachers. A list of consultants and documentation of their pertinent credentials is contained in the CLS files maintained by the Department (Appendix C). Clearly, the CLS staff, contractors, and consultants were well qualified to guide the developmental efforts before them.

In August, 1983, draft Competencies and Indicators, based largely on the work of Rosenholtz and Smylie, were formulated and mailed to state-wide Teacher Study Councils for review and comment. Responses from more than 6,000 teachers were analyzed and those indicators rated by at least 70% of the respondents as appropriate were retained. The resultant list was too exhaustive and unmanageable for measurement purposes.

To streamline the list, technical assistance was sought from Drs. Jason Millman of Cornell University and James Popham of UCLA; both are widely known and well respected experts on measurement and evaluation. The streamlined list contained six Competencies: (I) prepares for instruction effectively; (II) uses teaching strategies and procedures appropriate to the content, objectives, and learners; (III) uses evaluation to improve instruction; (IV) manages classroom activities effectively; (V) establishes and maintains a professional leadership role; and (VI) communicates effectively. A total of 18 Indicators were proposed as measures for the six Competencies. The final list of Competencies and Indicators appears to provide a justifiable representation of the research-based criteria proposed by Rosenholtz and Smylie and subsequently reviewed by more than 6,000 Tennessee teachers.

One Competency (#V) and two of its three Indicators are represented by the PDLS. The two Indicators are: (a) improves professional skills and knowledge and (b) takes a leadership role in improving education. A third Indicator, performs professional responsibilities efficiently, is represented on the Principal Questionnaire.

Completion, Submittal, and Scoring

The PDLS is the only written data source required by the CLS. It requires that the teacher candidate prepare a summary of professional development and leadership activities, without accompanying documentation, and submit it to the State evaluator on the first scheduled visit of the evaluation cycle.

The PDLS requires that teacher candidates complete and summarize activities in five areas for the professional development Indicator and in five or six areas for the leadership Indicator; six leadership activities are required for special populations and vocational education teachers to accommodate their particular needs. In order to "count," the beginning date of an activity may have occurred during any year of employment; but, activities must have been completed during the last five teaching years. Only one activity per activity area is permitted.

A copy of the current PDLS form appears in Appendix F. As shown in the Appendix, the activity areas included on the PDLS are as follows.

Professional Development Activity Areas: (1) obtains graduate degree(s) and/or takes courses; (2) participates in professional development activities; (3) uses ideas from professional books, journals, and professional organizations to enhance classroom instruction; (4) tries new methods/approaches in the classroom and evaluates success; and (5) seeks and utilizes community resources to enhance classroom instruction.

Leadership Activity Areas: (1) conducts workshops on training sessions for peers; (2) creates materials or programs and shares with peers; (3) holds a leadership position in the school/school system or educational organization; (4) promotes parent/community interest in the school; (5) initiates activities/projects in the school; and (6) communicates effectively with professionals, paraprofessionals, and/or parents (special populations only) or establishes and uses advisory groups (vocational education only).

For each activity, the candidate must summarize the amount of time spent on the activity, the activity's underlying rationale and professional goal, his/her role and responsibilities with regard to the activity, a description of the activity, benefits of the activity to the candidate's instructional role, and benefits of the activity to the candidate's students or to changes/improvements in the candidate's school/school system.

Each PDLS activity is rated on a five-point scale. The scale anchors are unsatisfactory, below average, average, outstanding, and distinguished. The ratings are averaged to generate professional development and leadership scores. As described in the Evaluator's Training Manual (e.g., Tennessee State Department of Education, 1988-89), evaluators are trained to rate the activities "wholistically." Nonetheless, they are also instructed to consider the following in determining the wholistic score: written statements of the rationale, purpose/goal, content, and scope of the activity, including the relative time and effort spent on the activity; the candidate's evaluation of the activity, its specific benefits, outcomes, and follow-up; and the degree of relationship between the activity and (a) the activity category as defined by the *Teacher Orientation Manual* (e.g., Tennessee State Department of Education, 1984-90), (b) the educational environment, and (c) the activity's application to the teaching setting. The candidate's ability to clearly communicate in writing cannot help but influence the evaluator's interpretation of the PDLS and, hence, the evaluator's ability to assign an accurate score.

Training

The basis for this section of the chapter includes participation in the PDLS evaluator training as well as extensive review of documents in the Tennessee Department of Education files. See Appendix C for a list of materials contained in the Department's files.

Not all evaluators are involved in scoring the PDLSs. Those who are receive approximately 12 hours of additional training. Training materials include an *Evaluator Training Manual* (e.g., Tennessee State Department of Education, 1988-89). Information in the *Teacher Orientation Manual* (Tennessee State Department of Education, 1984-90) parallels the *Evaluator Training Manual*. Although the Manual for evaluators is remarkably comprehensive, it is highly unlikely that it could ever replace the training itself.

In fact, the quality of the training is obviously dependent upon the trainer(s).

Data regarding the consistency of training across trainers are not available. However, it may be deduced that if the evaluators are proven reliable as a function of training, then the trainers are sufficiently consistent. Ultimately, nearly all evaluators demonstrate themselves reliable. As necessary, the training staff works one-to-one with evaluators who do not meet reliability criterion during the normal course of training. If criterion is still not met, the evaluator is assigned non-evaluation responsibilities.

The training consists of an introduction to the PDLS instrument, instruction in scoring procedures, and assessment of the reliability with which the evaluators assign PDLS scores. The goal of PDLS training is for the evaluators, by the end of the two-day training session, to meet reliability criterion of $\leq .80$ compared to ratings predetermined by experts (e.g., developers, staff). The reliability ratio is based on the sum of the differences across items of ratings assigned by the evaluator candidate (R) versus the criterion ratings assigned by experts (C), divided by the number of items being rated (N): $\frac{\Sigma(R-C)}{N}$.

Evaluator reliability is assessed by having the evaluators read the same sample PDLSs, already scored by a panel of "experts," and assign scores. Then, the extent to which the individual evaluator's scores match the scores assigned by the experts is determined. Expert raters are defined as those having expertise with regard to the content of the items being rated and in the structure of the instrument or rating system being used. In the case of the Tennessee CLS, the experts used to determine the standards for reliability were the training staff and the staff who worked on the development of the instruments. Reportedly, evaluators are regularly checked for reliability drift through reassessment at regional evaluator meetings (e.g., Malo, 1987).

For actual CLS evaluation purposes, each PDLS is independently scored by two evaluators. These evaluators then meet to reach consensus concerning their PDLS scores. In those cases where consensus cannot be reached, CLS staff arrange for a third evaluator to review the PDLSs in question.

Implementation

The procedures for implementing administration, submission and scoring of the PDLS are well documented in the Evaluator's Training Manual (e.g., Tennessee State Department of Education, 198-89) and the Teacher's Orientation Manual (Tennessee State Department of Education, 1984-90). Briefly, on the "A" or first visit of the evaluation cycle, the evaluator collects the PDLS from the teacher, fills out the appropriate receipts, keeps a candidate-signed receipt, and leaves a receipt with the candidate. Then, as described in the "Training" detection above, the PDLSs are submitted to CLS staff who assign them to two PDLS evaluators for independent reading. If scoring consensus cannot be reached, CLS staff consult with a third evaluator.

Data concerning "third party" or external evaluator observations of the extent to which the PDLS has been implemented as intended are not available. However, extant data indicate that for the most part, formative evaluation procedures utilized throughout CLS implementation (e.g., Sounding Boards, evaluator questionnaires) have been sensitive and responsive to the validity with which instruments and procedures, including those associated with the PDLS, have been implemented as intended. Such feedback led to the replacement of the Portfolio with the PDLS, clarification and refinement of the PDLS form without substantive alteration of the Competencies and Indicators, and a change in the PDLS submission date from the second to the first evaluator visit.

Summary

The Competencies and Indicators assessed by the PDLS have a research base (Rosenholtz & Smylie, 1983b), were rated as important by more than 6,000 Tennessee teachers, were retained through a consolidation process suggested by widely-known and well-respected technical experts, and approved by the Interim Certification Commission. Given the time available and the magnitude of the Competency identification and instrument development tasks, the content validity of the PDLS is reasonable, if not remarkable. Other aspects of its validity, namely the accuracy of the candidate's summaries and the extent to which the evaluator's ratings represent activity accomplishment versus writing ability, are more suspect.

The validity with which evaluators assign scores to PDLSs is necessarily dependent upon the reliability of the evaluators' ratings. The frequency and nature of evaluator drift checks is not clear based on extant data. Additionally, the validity of evaluator scoring is dependent upon the trainers and training procedures. Observation of evaluator training in the PDLS made it apparent that evaluator training is highly trainer-dependent. For example, although the ratings used as reliability standards must be consistent across training sites, the explanations or feedback given to evaluators may not be. The consistency of training warrants consideration. Also due further consideration is the match between intended and actual implementation.

Chapter IV: Content Validity—Dialogue Instrument

This chapter describes the development and implementation of the dialogue as one of the data sources used by the State of Tennessee's team of external reviewers to evaluate candidates for the upper levels of the Career Ladder System (CLS). It also discusses the development, implementation (including training) and relationship of the dialogues to the other data elements. It concludes with answers to the following four questions:

1. What evidence exists that instrument development procedures reflected the stated purposes of the CLS and reasonable practice?
2. What evidence exists that the instrument's content reflects specified needs, reasonable practice, and empirical findings?
3. To what extent do training procedures ensure that the instrument can be used as intended?
4. To what extent are the instruments and related procedures implemented as intended?

Instrument Development

The Dialogue is one of several CLS instruments used to differentiate among levels of teacher competency. A distinguished teacher reaches Career Level III, while the outstanding teacher reaches Career Level II. Teachers may apply for and receive Career Level I through local evaluation. Career Ladder II and III placement depends upon a complex scoring system (Rakow & Malo 1989). The Dialogues, which are interviews conducted by trained state evaluators with the Career Ladder candidate, are scored so that they may be aggregated with scores yielded by the other data sources: classroom observations, questionnaires and summaries of professional development and leadership activities.

Through a long, detailed process at the beginning of the CLS, it was decided that teachers should be evaluated for competency in five domains: planning, teaching strategies, evaluation of students and learning, classroom management and professional development and leadership. Of these five, the Dialogues focus on the first three domains.

In 1983, as the CLS was beginning, experts agreed that the competencies selected should be unequivocally related to student achievement. Despite an early decision not to use student achievement scores directly, a continuing theme was the relationship between excellent teaching and the improvement of test scores. It was decided that an interview by a state evaluator would be a direct way to collect data on the teacher's knowledge of planning, teaching strategies, and how to evaluate student learning. The assumption was that the potential for student learning was increased if the teachers could demonstrate knowledge in these areas. In fact, the final list of Competencies used as the basis for CLS instrument development was grounded in expert reviews of the research on effective teaching related to student achievement, validation by samples of Tennessee teachers, and final approval from the Interim Certification Commission which was composed of professional educators, including representatives of the Tennessee Education Association. A list of the Competencies underlying CLS instrument development is provided in Appendix G.

Because of the 1984 CLS authorizing legislation, State staff proceeded to develop instruments. This approach was necessary since appropriate instrument resources did not appear to be available (Malo & French, 1987). This customizing process led to a contract between the Tennessee State Department and the Appalachian Educational Laboratory to develop a Candidate Interview. Concurrently, a process to collect and evaluate candidate information via a Portfolio was initiated by State staff. Eventually, these two data bases were reduced and merged.

The interview instruments and processes for administering and scoring them, were subsequently piloted statewide using a random sample of schools. Then, during 1984-85, 100 teachers evaluated 3,350 of their peers for either Level II or III of the Career Ladder.

The first state evaluators were selected from among applicants having at least eight years of experience. Preference was given to those who had teaching experience in more than one grade level or subject matter area. A selection panel screened the qualified applicants and then contacted their local school systems to confirm availability and obtain recommendations. These evaluators were and remain the heart of the CLS.

Rather than having State staff as evaluators , the CLS legislation originally required that a three-member team of peer evaluators from outside the candidate's own system be trained to administer a common set of instruments, using standardized procedures to collect data and generate scores. These peers would agree to spend a year of their time to become trained and, once trained, to travel throughout the state conducting the evaluations. As the System matured, only those teachers who had gone through the process would be eligible to be peer evaluators. Using this approach neither State bureaucrats nor a teacher in the same school or district as the candidate would be involved in a decision that not only means status, but also a great deal of money over a teacher's career.

A planning grant submitted by the Tennessee State Department to the U.S. Office of Education (1986) contained several comments and recommendations about the interview development phase of the CLS. While the first three recommendations are truisms about almost any state program (not enough time for field testing, the problems of validity when customizing an initial instrument, and the need for a staff person to document the development of the system and to organize records), the report also contains a section that sheds light on Tennessee's particular interview process. Apparently due to time constraints, "there was not enough communication between the State Department and the contractor. When the contractor arrived at the field-test training, their interview was too long and had to be reduced from

32 question sets to 14. Not only was the administration time of the instrument too long, but problems also existed with the training program."

Because the contractor was present for interview-related sessions only, the philosophy, style, and presentation across sessions were not consistent. In addition, not nearly enough time (two weeks) was allocated for the entire training effort. Although statistical procedures to assure interrater reliability were in place, the timeframe was insufficient to train the evaluators in the Observation Instrument, Interviews, Portfolios, and Questionnaires.

The first-year Portfolios were a paperwork nightmare for teachers since they were required to develop instructional plans, classroom management procedures, evaluation procedures, and leadership and professional development activities for inclusion. According to CLS staff (Furtwengler 1987), teachers spent endless hours on their Portfolios--an expenditure of time that did little to enhance classroom instruction. Our review of the Teacher Portfolio Rating and Summary document supports the concerns of candidates, staff, and evaluators from the 1984-85 experience regarding data burden for all parties.

The first-year Interviews also required extensive documentation, partially to prepare for any appeals. The State evaluators had to document the candidates' answers to each of the Interview questions. These Interviews were structured, controlled by the evaluator, lasted several hours, and proved to be a time-consuming ordeal for the evaluator and a stressful and exhaustive experience for the teacher (Furtwengler, 1987).

The law required a comprehensive, in-depth evaluation system that would be fair and objective and that had high standards for identifying outstanding teaching performance. Therefore, the initial, complex attempts to establish multiple data bases were due to be refined after the first year. In fact, the Portfolio was eliminated in the second year, except for the candidate's report of professional development and leadership activities. The Interview was replaced with three one-hour "Dialogues" with the candidate, one in each of three areas: Planning, Teaching Strategies, and Evaluation. Each Dialogue was conducted by a different evaluator. During each Dialogue, 30 minutes would be devoted to structured questions and 30 minutes to unstructured time in which the teacher shared information informally with the evaluator (Millman 1985).

The initial version of the Interview took more than three hours to conduct. In replacing the Interview with the three Dialogues, entire question sets were deleted from the process. This procedure was used so that the number of questions per rating, and therefore, reliability of each rating, would not be jeopardized. The question sets retained by the Dialogues were selected because they were not as well covered by alternate data sources.

For the 1985-86 administration of the CLS, only the Leadership Summary section of the original Portfolio remained. It became part of a new instrument, the Professional Development and Leadership Summary. The Dialogues were revised generally to the present form:

- One Dialogue occurs during each of the three visits.
- Each Dialogue has a one-hour time limit.
- Structured and unstructured time is built in to each Dialogue.
- The Dialogues separately focus on three domains: Planning, Teaching Strategies, and Evaluation.
- The Planning Dialogue contains three question sets: (1) developing goals/objectives, (2) determining how to teach and use instructional materials, and (3) accounting for student differences in all phases of the classroom plan.
- The Teaching Strategies Dialogue contains four question sets: (1) describing how the learning will take place, (2) monitoring student learning, (3) providing practice and review opportunities, and (4) keeping students involved.
- The Evaluation Dialogue contains three question sets: (1) evaluating student progress, (2) using evaluation results to modify teaching, and (3) developing as well as communicating high student academic expectations.

Modifications over the years, up to and including the training of evaluators for the 1989-90 school year, concentrated on fine tuning the questions asked by the evaluators and giving the interviewer more flexibility to ask probing questions. A copy of the Dialogue, in its present form, appears in Appendix F.

Teacher Orientation to the Dialogues

The Tennessee State Department annually publishes a Career Ladder Teacher Orientation Manual (Tennessee State Department of Education, 1984-90) which describes in some detail the entire program, including the actual instruments used by the peer evaluators during the three dialogues. This manual is designed to provide candidates with the information necessary to understand and successfully participate in the CLS.

A review of the Teacher Evaluation Manuals from 1985 to 1989 shows very few changes in the Dialogues or the associated criteria used to evaluate the candidate. The 1989 version does have a few additions to help explain the process. For example, the candidates are informed that the teacher may wish to use notes, but should not attempt to read a script; that the evaluator may interrupt to ask questions; that the evaluator will be concerned more with why a candidate does something than with what they do; and that any material the candidate brings to the Dialogue is not rated by the evaluator, but that it may help the evaluator in scoring the dialogue.

The Planning Dialogue

The Teacher Orientation Manuals (Tennessee State Department of Education, 1984-90) are designed to focus the candidate's actions during each one of the three meetings. In preparing for the first

Dialogue, which focuses on the Planning Domain, the candidate is told to prepare to discuss:

1. How they decide what to teach. They are to briefly outline their instructional program for the year, discussing the goals and objectives of the unit and lesson plan that was observed. It is quite clear that at least one classroom observation must occur prior to the planning dialogue.
2. How they use teaching strategies and instructional materials.
3. How they adjust their plans according to student differences.

These overall questions are then followed by an array of tips which the State believes will be helpful to the candidates. The candidates are warned that the evaluator will take notes but will not record the interview verbatim and that the evaluator completes the scoring after the visit is over. This allows the evaluator time to give careful consideration to all the responses before assigning a score.

The above three question areas drive the focused half of the Planning Dialogue. During the remaining 30 minutes, which constitute the unstructured portion of the Dialogue, the candidate has an opportunity to share with the evaluator those things the candidate believes are important. Even though this sharing is unstructured, it is important because the information gathered there will be part of the basis for the evaluator's score.

The Teaching Strategies Dialogue

The candidate/evaluator dialogue during the second visit focuses on the Teaching Strategies Domain. The candidates should be able to discuss:

1. How they explain their content or subject area to their students, that is, what the students are to learn and how it will be accomplished.
2. How they keep track of student learning during a lesson and what they do when some students do not understand the lesson.
3. How they provide for student practice and review after a lesson has been introduced.
4. How they keep students involved and participating in learning activities.

The candidates are encouraged to share information about their effective teaching strategies and to bring materials which support their answers to the above questions. During the Teaching Strategies Dialogue, the evaluator also rates the teacher on the use of grammar and oral communication skills.

The Evaluation Dialogue

The third visit focuses on the Evaluation Domain. Thirty minutes of the dialogue are devoted to the following questions:

1. How the candidates evaluate what they teach, especially as shown through student progress.
2. How they use their evaluation results to enhance student learning.

3. How they develop their expectations for the students and communicate those expectations.

In other words, the teachers are asked to explain what they intend to teach; how they intend to teach it; how they know if the students learned the material; and if they didn't learn it, how they will reteach the material for these students.

Scoring

In addition, the Teacher Orientation Manuals (Tennessee State Department of Education, 1984-90) describe precisely how each of the three Domains of competence (Planning for Instruction, Teaching Strategies, and Evaluation of Instruction and Student Progress) are evaluated by the Dialogues and how scores are assigned. Each Domain has a series of Indicators or specific descriptors. These Indicators appear at the beginning of the Manual and hold for all the evaluation procedures. However, within the Dialogue they are reworded and used as the basis for the scoring. For example, in the Planning Domain, competency A "Establishes appropriate instructional goals and related objectives consistent with the curriculum" is assessed using four Indicators, each of which is rated on a five-point scale ranging from unsatisfactory to distinguished. Each of 11 items is similarly assessed such that scores may range anywhere from 11 to 55 points. The same pattern holds for the other two Domains in the remaining two visits. The rating statements and protocols for the dialogue instrument are shown in Appendix F.

The Tennessee State Department publishes a Career Ladder Technical Manual (e.g., 1988-89) as a guide for teachers as they interpret their scores. Candidates receive as part of their printout an average score for each Indicator within each Domain. For example, the four scores (from 1 to 5) for Indicator A in the Planning Domain are totaled and divided by 4. The most recent means available show that the average score for Indicator A was 3.889 with a standard deviation of .712. This means that 68% of the candidates' scores fell between 4.601 and 3.167. The typical rating was close to 4 which is equivalent to "outstanding" with the vast majority of the candidates between "average" and "distinguished."

Only the Indicator of Evaluation Focus, within the Evaluation Dialogue, shows a difference in average scores compared to the remaining 10 Indicators. That is, scores dropped when the candidates were asked to explain how they establish and communicate their expectations for student learning, group their pupils, document academic achievement commensurate with background ability of students, and show improvement in student attitude toward learning. The average score for this Indicator is 2.61 with 68% of the scores falling between 1.7 (below average) and 3.705. The Career Ladder Technical Manual (e.g., Tennessee State Department of Education, 1988-89), therefore, provides a step-by-step way for the candidates to understand how their scores were developed for each instrument, by domain, and by indicator. It also explains the appropriate weighting and placement of scores on a standard 200-800 scale and provides a State benchmark for candidates to use in understanding their scores in relation to their peers across Tennessee.

The above discussion has centered on the Dialogue Instrument and procedure, only one of several data bases used to generate a score which determines whether a candidate becomes a CLS Level II or III teacher. The philosophy of the program has been built on the principle that no one instrument or procedure is so comprehensive that a reasonable evaluation decision can be made by its implementation alone. Thus, multiple data bases are necessary. In addition, since it was decided at the beginning that a student test would not play a role in making these decisions, a labor-intensive approach using teams of objective peers seemed desirable and necessary. To ensure that all candidates would be treated equally and fairly required that these peer evaluators undergo an intensive training program.

Evaluator Training

Beginning each August, all evaluators complete a month of initial training in various phases of the evaluation process. A number of days throughout the year are spent in additional training and refining skills. According to CLS staff about 225 hours annually are necessary (Malo 1987.) About 50 of these hours are spent learning to conduct the Dialogues. Evaluators are tested for reliability and cannot actually conduct the Dialogues until they have passed pre-established reliability tests, that is, assurances that evaluators can show consistency in their ratings. Evaluators are trained to use the instruments and procedures in all subject matter areas and grade levels, not just their own areas of certification and experience.

The law enables the state to borrow Career Level III teachers from local school systems for a full academic year. The State pays the salaries of these full-time evaluators thereby freeing local funds for the salaries of substitutes.

The following comments are based on our 1989-90 observation of a week of evaluator training to conduct Dialogues. The training was conducted by a team composed of staff from the Department of Education's CLS Office and by individuals who had been peer evaluators in previous years.

The evaluators are taught how to be good interviewers and good listeners. They are to suspend their personal definition of good teaching in order to establish a common basis for consistent rating across the state. Interspersed with the general training are specific sessions in understanding the Domains and Indicators and in scoring procedures.

Key observations during a week's worth of Dialogue training include the following:

- Evaluators are to keep the teachers focused during each Dialogue and not let them wander. A good interviewer can be an unfocused candidate's best friend. For example, during the Planning Dialogue, the interviewer tries to keep the candidate from discussing teaching strategies until the next visit.

- Interviewers can only ask questions for clarification. They are specifically taught the difference between leading and clarifying questions. For example, evaluators are told that it is not sufficient to accept the statement: "I use the state curriculum in my planning." Interviewers must ask how the teacher candidate operationalizes the state curriculum guidelines. The interviewers must guard against inferring from their own experiences what the answer ought to be.
- Among the points of philosophy the evaluators learn is that all good teachers use grouping in some way; hence, a candidate's statement in the evaluation domain that "I have a homogeneous class so I don't group" receives a low score. Teachers can group by ability, achievement, interest, learning style, subject area, age, etc. If the candidate does not mention grouping, then the evaluator is not to bring it up during the interview. The interviewer cannot lead the candidate to the topic. The Tennessee system is so open that extra help is not deemed to be necessary. For example, the Domains, Indicators, Dialogue Questions, and Criteria are provided in documents such as the Teacher Orientation

Manual (e.g., Tennessee State Department of Education, 1984-1990).

- Evaluator's questions are to clarify, focus, probe, expand, and add depth. Some of these questions may well relate to the materials which the candidate presents five minutes prior to the Dialogue. This is a leftover from the previous Portfolio concept. The materials are supposed to provide a useful background for the interviewers. Therefore, they are told to seriously read and review them. These materials can be the basis for probing questions during the Dialogue. Evaluators are told not to take the materials with them and to assure the candidates that the materials themselves are not evaluated. The issue of how thorough the interviewers are in this review is presently being attacked through an appeals process. Some candidates not attaining their desired level say that the interviewers did not really review their work or discuss it during the course of the interview. The interviewers are told to stay within the five-minute guideline. If the time varies from one candidate to another, it is appealable by a losing candidate.
- Distinguished teaching occurs independently from a variety of factors. For example, regardless of the teacher's statement about the constraints of working

within state or district guidelines, it is the teacher's classroom interpretation that makes the difference.

- Evaluators are trained to review what the candidates say according to the criteria specified in each Indicator's measurement statements. On the 5-point scale used to evaluate responses, 3 is typical, and in fact, the interviewers are told that it is their reference point. A response that is less than average is scored a 2 unless it is absent or a low level and then is scored a 1. The same procedure is used going up the scale in moving from above average to distinguished. However, the evaluators are told that just because a teacher receives a 4 doesn't mean that the teacher should be at CLS Level II and if a teacher receives a 5 that the teacher is necessarily ready for Level III.
- The State staff emphasize the importance of listening skills and being open, sensitive and attentive to the candidates. For example, if the evaluator does not ask questions, the candidate may assume that the evaluator is not listening or "just doesn't care."
- A major purpose of the Dialogue is to obtain information that cannot be captured during the classroom observations. The classroom observations do not provide direct systematic evidence about teacher planning, teacher assessment and modification of instructional materials, and teacher choice and adaptation of

instructional materials (Stodolsky 1990). However, at least one classroom observation has to occur prior to the Dialogue. The evaluators are warned about the halo effect, that is, assuming the planning is good just because the observation is good. Evaluators are drilled on the concept that good teaching is good teaching no matter where the classroom is located. Just because a teacher does not have enough money to buy all the soccer balls he/she wants and has to work in a poor facility does not mean that the essentials of good planning cannot exist. Also, higher order thinking skills can be incorporated in physical education as well as physics. Through the use of simulations and video tapes of Dialogues, the evaluators also were trained not to make judgmental comments or interject their opinions into the Dialogue. The tapes are also the basis for starting the process of determining each evaluator's ability to score the Dialogues according to the predetermined scores of expert raters. There is a great deal of group viewing of tapes, evaluation, debriefing, and then feedback to operationalize all the measurement statements that are listed in each Domain. Evaluators are continually pressured to adopt the criteria and standards of the CLS at the expense of setting aside their own standards. This is similar to approaches used to train scorers of essays in which all of the papers must be evaluated according to common criteria to assure that the final scores are fair and objective. To accomplish this, the evaluators are trained to think wholistically and not in a checklist fashion.

- Finally, all the evaluators are told that when unsure about a score to always give the benefit of the doubt to the candidate by assigning the higher score. Fairness and objectivity were stressed in all the sessions.

Discussion

The Career Ladder Teacher Orientation Manual distributed to teachers across Tennessee (Tennessee State Department of Education, 1989-90) clearly states that the two primary goals are to identify and reward outstanding teaching performance and to improve instruction. Further it states that the CLS is based on two premises: 1) a sound evaluation program focuses on performance, rather than credentials and 2) it is possible to assess differences in teacher performance.

Given the above statements as the standard to use in validating the existing Dialogue for collecting information, the four questions cited earlier will be answered in this discussion section:

1. What evidence exists that instrument development procedures reflected the stated purposes of the CLS and reasonable practice?

2. What evidence exists that the instruments' content reflects specified needs, reasonable practice, and empirical findings?
3. To what extent do training procedures ensure that the instruments can be reliably used as intended?
4. To what extent are the instruments and related procedures implemented as intended?

Question #1

The CLS is characterized by a multi-instrument approach to collect objective, reliable information to ascertain who is an outstanding teacher. Instead of relying on one instrument or procedure, the CLS is built on the concept that no one test or procedure can be used and accepted as having the ability to discriminate among the very good and the outstanding. Therefore, a number of different instruments must be used to present a comprehensive evaluation. The Dialogues represent a key component of the System since they allow teachers to fully describe how their teaching meets the standards in three of the key identified Domains: Planning, Teaching Strategies, and Evaluation of students and learning.

Available documentation (Appendix C) shows that the State used recognized procedures in obtaining the best advice possible in identifying the Domains and related Competencies. Tennessee teachers had a chance to comment on them, and the Tennessee Education Association provided continuing advice. National experts were brought in from the beginning and their advice was sought throughout the developmental stages. The consensus from the experts and the review of the literature led to the justified conclusion that instruments and procedures would have to be customized to meet the intent of the Tennessee law and the Certification Commission established by the law.

Short timelines led to the initial over-development of Portfolios and structured interviews. These initial efforts were an attempt to collect all possible information, but were burdensome to the point that the System would likely have collapsed internally from the weight of the data and the cries of the teachers about the paperwork. In retrospect, it was unwise to have undertaken large-scale use of Portfolios for school teachers in the virtual absence of small-scale experience (Bird 1990). But the staff modified the System, again with the help of outside experts as well as the Tennessee teachers who participated during that first year, to focus it on the three key Domains now included in the Dialogues and to make it manageable.

Since that first year, the Dialogue system has remained fairly constant. The questions, including the subpoints under each question that are used to determine point totals, and even tips for the candidates' consideration as they prepare for the visits are all open and available for review. The instruments, their purpose, and relationship to the Domains are clearly described.

The original legislation contained quotas by school system in the percent of teachers who could enter the top levels of the CLS. However, the final bill did not put a percentage limit on the number of teachers who can qualify at the upper levels, nor did it limit the location of the teachers by school system (Furtwengler 1985). The Dialogue Instrument is just one example of how all who qualify will be rewarded--that teachers compete against an external standard instead of each other (Glass 1990).

In summary, the statistical analyses of results provided in Chapter VI indicate that the Dialogues are critical in deciding the final Planning, Strategy, and Evaluation Domain scores. The Dialogue section of CLS contains reasonable and valid instruments and procedures to select distinguished teachers. Candidates who are eligible to apply for the Career Ladder II and III positions have access to all the information necessary to prepare for the Dialogue and to make a case for their claim to be on the top rung of the ladder.

Question #2

With regard to the second question about the content of the Dialogues, our primary conclusion is that the questions are fair to all teachers across grades and subject matter areas. For example, they can apply equally to a high school physics teacher, as well as a junior high physical education teacher or a primary self-contained classroom teacher. In fact, the evaluators are drilled during their training procedure that good teaching in all these areas must have the same high levels of planning, strategizing, and evaluating. We have no evidence that the questions nor procedures are changed according to whether the teacher is from a wealthy or a poor school or is black or white.

Our examination of the extant data indicates that the content of these questions has remained from the earliest days of the expert and professional review.

One of the key alterations the first year was to change from a formal interview where the teachers had to respond to an extensive list of questions to the present format where the evaluator provides a template via a few questions and the teachers tell their story of how they plan, actually teach, and evaluate whether the pupils have gained the desired knowledge. Although the data we reviewed did not formally describe why the name of the procedure was changed to "Dialogue," it certainly was appropriate. During the training, the evaluators are told that this time is the candidates' opportunity to explain their philosophy and how they turn their theories into action. The evaluator simply tries to keep the conversation focused.

Question #3

Valid instrumentation and procedures are necessary but not sufficient unless the evaluators are trained to use them objectively, fairly, and reliably. The materials, our observations of evaluator training, and interviews with evaluators and trainers have provided ample evidence of the high quality of the System. Training any large group to hold their own perspectives in abeyance while they learn a common definition is always difficult.

The most essential element in a training program, especially one that goes on for a month, is a dedicated, knowledgeable staff with sufficient materials. The training that we observed this year had not only several State staff members who have been with the CLS since its beginning but also several former evaluators. Their formal and informal assistance was quite valuable to a group of strangers from schools all over the state who had left the safe environment of the classroom to become evaluators of their peers. The smoothness of this process is greatly dependent upon the accumulated experiences of the staff.

Much of the evaluators' positive self-concept and confidence in their own abilities occurs as a result of the statistical checks used to evaluate progress. For example, during the week devoted to training on the Dialogues, the evaluators viewed video tapes of interviews. Eventually, the staff felt the group had reached a point where they could rate the tapes using the Dialogue Instruments. Their scores would be compared with those previously assigned by a panel of experts. This procedure enabled the staff to evaluate the progress of the group as well as individuals and to plan future training sessions. Reliability checks to evaluate consistency and accuracy continued throughout the training and also in the field. Staff used a standard statistical formula to document reliability before moving to the next activity. Staff had procedures to check for rater drift, that is, when an evaluator started to systematically score a candidate higher or lower on the questions compared to the criterion scores set by experts.

Training procedures are in place to ensure that the instruments are reliably used as intended. However, it is recommended that a portable or lap-top computer along with a printer should always be available for CLS staff in conducting the reliability checks. Also, additional video tapes should be made available for training purposes. Perhaps, with the emergence of video discs, a training package could be put together to enable some regional training.

Question #4

Finally, the Dialogues are valid elements of the CLS in that they meet the letter and spirit of the law. They are well founded in research and statistical checks for reliability meet accepted practices.

Analysis

The previous section discussed the CLS according to the questions of validity as they relate to the intent of the law. However, unanticipated consequences of the program should also be explored and the original purpose of the law should be reviewed.

Without a doubt, Tennessee has an extraordinary resource that should be the envy of every other state: the cadre of evaluators who have undergone a full year of training and who have observed the best of their peers throughout the state. In a recent dissertation, Parkison (1987) describes the reactions of the evaluators whom she interviewed and surveyed. These evaluators said that when they returned to the classroom they perceived themselves as having improved their planning for instruction, strategies on how to teach, and evaluation skills. They attributed these improvements to their exposure to the ideas, methods, materials, procedures, and teaching styles of other teachers, as well as to their training on how to systematically analyze a teacher's behavior. The evaluators also reported a significant gain in personal growth and attributed this change primarily to the independence, self- reliance, and decision-making demands of being an evaluator and participating in evaluator training.

Obviously, hundreds of Tennessee teachers believe they have undergone an extraordinary amount of professional growth that would not have been possible without their participation as evaluators in the CLS. A question that Parkinson brings up for the State and its institutions of higher education is how can an even larger segment of Tennessee teachers benefit from the CLS training procedures in order to analyze and improve their own teaching?

While the observed quality of the training is superb, it is too person-dependent at this time. That is, the staff in charge have not had an opportunity or funds available to develop a polished package that could be exported. Staff seem to have a sixth-sense about when to adjust a sequence of materials, change direction or skip ahead based on their experience in working with large numbers of trainees. These organizational guidelines are not in writing, nor is there a polished package of written or audio-visual materials. Expansion to regional training for CLS evaluators is out of the question until this package is together; the credibility of the CLS depends on teachers across the state knowing that the quality of the training is consistent and does not vary by where the evaluator received the training or who happened to staff a certain week of training.

The proceedings from the 1988 Educational Testing Service (ETS) Invitational Conference, entitled New Directions for Teacher Assessment, provides another standard from which to view the progress of the CLS. In that document, Kelly (1989), as part of his argument for establishing a National Board for Professional Teaching Standards, states that "though dozens of accountability measures have been spawned, many educators, as well as government and business leaders, continue to express doubts concerning the quality of teaching in schools." These statements are similar to those made five years earlier in the Tennessee debate over the CLS. A further complaint made by Kelly is the fact that, "little distinction is made between the first-year and twentieth-year teacher."

Kelly goes on to state that the National Board for Professional Teaching Standards is a non-profit organization whose main purpose is to establish high and rigorous standards for what teachers should know and be able to do and to identify those teachers who meet the standards. The proposed standards sound very much like the Indicators within the Domains assessed by the CLS: teachers should treat all students equitably and understand the personal and instructional implications of student differences; teachers should be skilled at diagnosing and facilitating learning; armed with a variety of methods and modes of interaction, they must create, enrich, and on occasion alter the organizational structures in which they work with young people; and they must sustain the interest of students and evaluate their students' learning.

In the same document, Shulman (1989) adds that teacher assessment must measure teacher classroom management and organization skills as well as what, how, and why teachers think about their actions in teaching particular ideas and skills. Shulman goes on to argue that "teaching is such a complex and contextualized phenomenon that any single mode of measurement will fail to assess its practitioners validly...the solution lies in deploying complementary modes of evaluation." Therefore, he, as well as others pursuing a National Board, advocate the Tennessee philosophy of a multiple data base for making decisions. Interestingly enough, Shulman argues for the use of portfolios—a concept that was tried and rejected in Tennessee.

In these same proceedings, Urbanski (1989) praises the Rochester, New York experiment because "unlike merit-pay systems that purport to be career-ladder programs, our Career in Teaching plan incorporates the peer-review concept." Tennessee made a significant decision at the beginning of its efforts to avoid the easy solution of making evaluators full-time career positions instead of annually selecting and training a new contingent of teachers. But the decision has gained credibility since the evaluation process is peer-based, and as mentioned at the beginning of this section, has produced a statewide core of highly trained classroom teachers who can be school, district, and regional resources.

Perhaps, the most significant question regarding the validity of the System is not whether the pieces are the right ones and whether they mesh to really identify the most distinguished teachers, but rather should the system have been built at all? Timpane (1989) concludes the ETS conference proceedings by stating that:

> We can have no doubt that the movement toward teacher assessment in this nation is substantial, durable, and nationwide...for many years we thought the testing of students would establish this accountability, and when that didn't work, we tried the evaluation of programs, and when that didn't work, we began to zero in on the performance of the individual educator...the quest...is a growing feature of our society, not simply of our educational systems ...there is nothing in the policy world that suggests this direction toward greater accountability and more substantial teacher assessment is going to change...the pace may change; the details may change. Some ideas and directions may prove to be worthwhile, some not. But we will not, as a nation, retreat from this question. It is a part of the larger quid pro quo that seems to govern the educational reform movement today--that if there are going to be more resources for the schools and better pay--the quid--then the policy makers of the world are going to insist on a quo.

Timpane goes on to ask whether teacher assessment will improve the education of children and quite appropriately answers that it depends. It depends on whether or not the assessment is designed and carried out in an organized and systematic perspective in league with other policies for school improvement, better teacher preparation and improved teacher-induction programs. Perhaps, another way to phrase the larger question not addressed in this study is whether the CLS is viewed, in fact, as one of many interrelated means to the end of improved pupil performance or as an end itself.

There is one final note regarding emerging methods of teacher assessment that may have implications for the CLS. Haertel (1990) writes from his experience at Stanford University's Teacher Assessment Project that performance testing and simulation lie somewhere between observing teachers in their classrooms and multiple-choice tests. He goes on to say that these methods are only now being explored with prototypes under development, but he believes they promise to assess a greater range of teaching knowledge and skills with greater reliability and efficiency than classroom observations. Among the performance exercises now being developed are one requiring teachers to present a lesson to a small group of students, another in which a teacher examinee views and discusses a video tape of another teacher's performance, and yet another in which a teacher responds to students' scripted questions.

Summary. Through the 90's, emerging, innovative assessment approaches, such as those described above, should be monitored to see if they could supplement the present Dialogue and other instruments. But at this time, the CLS Dialogue is working as intended.

Chapter V: Content Validity—Questionnaires

The Elementary Student (Form E), Secondary Student (Form S), and Principal Questionnaires are used along with other Tennessee Career Ladder System (CLS) instruments for making decisions about individual teacher quality. The other instruments are described in previous chapters and include the Dialogue, Observation Instrument, and Professional Development and Leadership Summary (PDLS). The total CLS score, upon which CLS advancement decisions are made, is based on a weighted sum of scores on these instruments plus a Professional Skills Test.

The purpose of this chapter is to examine the content of the Questionnaires in order to determine the extent to which they can be expected to contribute to accurate CLS advancement decisions at Levels II ("better") and III ("best"). The questions driving this chapter are the same as those that guided the preceding chapters on the content validity of the Dialogue, Observation, and PDLS instruments:

- What evidence exists that instrument development procedures reflected the stated purposes of the CLS and reasonable practice?
- What evidence exists that the instrument's content reflects specified needs, reasonable, practice, and empirical findings?
- To what extent do training procedures ensure that the instruments can be reliably used as intended?

- To what extent are the instruments and related procedures implemented as intended?

Principal Questionnaire

Instrument Development and Content

Purpose and Chronology of Development/Implementation. The Principal Questionnaire provides an alternate or additional data source for each of the CLS Indicators. The Questionnaire consists of items designed to assess Indicators of each of the Competency areas assessed by the CLS: (I) prepares for instruction effectively; (II) uses teaching strategies and procedures appropriate to the content, objectives, and learners; (III) uses evaluation to improve instruction; (IV) manages classroom activities effectively; and (V) establishes and maintains a professional leadership role. Together, the Competency areas encompass a total of 17 Indicators for General Education; two additional Indicators (total n=19) apply to teachers of special populations and one additional Indicator (total n=18) applies to vocational educators. The additional Indicators are included to meet the particular evaluation needs of special populations and vocational education candidates. Hence, the Principal Questionnaire contains 20 items or Indicators, with a maximum of 19 Indicators applying to any one candidate.

A review of documents on file with the Tennessee State Department of Education (Appendix C) indicates that the Principal Questionnaire was first used in 1985-86 during the second year of

CLS implementation. A "Superordinate" Questionnaire, used in 1984-85, was the precursor to the Principal Questionnaire. The Superordinate Questionnaire, completed by the person responsible for the candidate's personnel evaluation--usually the Principal--consisted of 45 items, provided no explanations regarding the various items, and used a frequency scale which asked how often a teacher did something (i.e., half-of-the-time, usually, almost always). The subsequent Principal Questionnaire consisted of one item per Indicator, provided a principals' handbook containing explanations pertaining to the items and their completion along with an explanation of the principal's role in the evaluation process, and used a quality scale asking how well the candidate did something (i.e., average, outstanding, distinguished).

The decision to replace the Superordinate Questionnaire with the Principal Questionnaire is documented in the minutes of the 8/16/85 meeting of the Interim Certification Commission. The decision was based, in part, on technical advice from Dr. Jason Millman, a measurement expert and Professor of Education at Cornell University, who observed the evaluator consensus judgment process, inspected the CLS instruments, guided the design of data analysis of the first year's evaluation results, and studied the first available results of the data analyses. With regard to the Superordinate Questionnaire, he advised that the scale be changed from one of frequency to one of quality, that the number of questions be reduced, and that the evaluator ask for examples of candidate qualities that the superordinate judged as outstanding.

Other than wording refinements and improvements in instructions and explanations there were no changes to the Principal Questionnaire in 1986-87. In 1987-88 items were field tested for inclusion in the 1988-89 instrument. A copy of the instrument now in use, including the Competency and Indicator associated with each item, appears in Appendix F. The changes in the instrument since its inception as the "Superordinate" Questionnaire seem to have added needed clarity and increased the potential for both implementation and scoring validity.

Competencies and Indicators. As with the other CLS instruments, the identification and content validation of Competencies and accompanying Indicators for CLS inclusion was based on effective teaching research, consensus by practicing teachers, and expert opinion. In no case does the Principal Questionnaire provide the only data source for an Indicator. Rather, it provides another data source per Indicator, based upon a particular perspective, i.e., that of an administrative supervisor. Refer to Appendix G for a list of the instruments which address each of the Indicators included by the CLS for general education, special populations, and vocational education candidates.

The Competencies and Indicators were approved by an Ad Hoc Interim Certification Commission convened in April, 1983 and charged with developing a master teacher evaluation system. Staff to the Commission included an executive director, Dr. Russell French from the University of Tennessee at Knoxville, and personnel from the State Department of Education's Division of Research and Development.

Vanderbilt University was contracted to provide research information on teacher effectiveness and indicators of teacher competencies which affect student achievement. This research formed the bases for the content of the various CLS evaluation instruments. The resulting Competencies and Indicators were submitted to review by more than 6,000 teachers.

As stated previously, all of the Competencies and their associated Indicators are represented on the Principal Questionnaire. Each item on the questionnaire addresses one indicator. The indicators are included in Appendix G.

Administration of the Principal Questionnaire. During the first evaluator visit of a particular evaluation cycle, the Evaluator ensures that the principal has a handbook that describes the evaluation process and explains the Principal Questionnaire. The evaluator also leaves a copy of the Questionnaire with the principal for collection by an evaluator at the second visit.

The completed Principal Questionnaire is collected by the evaluator conducting the second visit. Completed questionnaires are submitted to CLS staff for scanning/scoring and inclusion in the final evaluation score/report. Principals may share their responses on the Questionnaire with the candidate if they wish. Regardless of the principal's decision in this matter, the candidate will see the scores from the Principal Questionnaire on the final evaluation report. This practice may influence principals' responses to the Questionnaires.

Training/Implementation

This section of the chapter is reliant upon extant information available in the files outlined in Appendix C. Training in questionnaire related procedures was not directly observed by members of the validation team. According to a technical paper presented by Malo (1987), training in questionnaire administration is approximately four hours total including the Student Questionnaires.

A review of training materials (e.g., Evaluator Training Manual, 1988-89) indicates that most of the instruction centers on the logistics of questionnaire administration. In the case of the Principal Questionnaire, instructions are quite straightforward. During the first visit, the evaluator makes sure that the principal has a handbook which describes the principal's role in the CLS evaluation process and provides guidelines and explanations pertaining to questionnaire completion. The extent to which principals adhere to the guidelines is not easily assessable. The files (e.g., Sounding Board summaries, ICC minutes) indicate that modifications in the Instrument over the years, such as item clarification via handbook explanations and a reworking of the format, were targeted at increasing the extent to which principals used the Instrument to differentiate between "better," and "best" candidates. Principals tended to rate toward the high or "best" end of the scale. The most recent assessment of the reliability of the Principal Questionnaire provides weighted alpha coefficients of .78-.95 for five domains as follows: Planning (.94), Strategies (.95), Evaluation (.81), Management (.88), and Leadership (.78).

Student Questionnaires

Instrument Development and Content

The Student Questionnaires are designed to gather student perspectives on the Indicators listed below. The underlying assumption is that students are in a unique position to observe Indicators of these Competencies and judge them.

Competency II (uses teaching strategies and procedures appropriate to the content, objectives, and learners):

- Provides a clear description of the learning task and its content.
- Monitors learner understanding and reteaches as necessary.
- Provides learners appropriate practice and review.
- Establishes and maintains learner involvement in the learning task.

Competency III (uses evaluation to improve instruction):

- Reports learner status and progress to learners and their parents.
- Improves learner performance.

Competency IV (manages classroom activities effectively):

- Establishes and maintains appropriate learner behavior.

- Establishes and maintains a classroom climate conducive to learning.
- Makes effective use of classroom resources.

The Elementary and Secondary Student Questionnaires appear in Appendix F. The Competency and Indicator addressed by each questionnaire item is provided. The Secondary Questionnaire contains 39 items for response by all students with three additional items for response by vocational education students whose learning situations may be somewhat unique. The Elementary Questionnaire contains 34 items. Insofar as appropriate, the Elementary and Secondary items parallel one another. However, wording or number of items per Indicator varies by Questionnaire so that the Questionnaires may be optimally sensitive to the different experiences and levels of understanding of the elementary and secondary school students.

A review of documents on file with the Tennessee State Department of Education (Appendix C) indicates that the Student Questionnaires were first used during the first year of CLS implementation in 1984-85. At that time, there were three questionnaires, one for grades K-2, 3-6 (Elementary form), and 7-12 (Secondary form) consisting of 9, 34, and 44 items, respectively.

In 1985-86, alternative administration methods were allowed for the K-2 Questionnaire to accommodate the needs of primary children (e.g., marking directly on the questionnaire instead of on a computer-scanable answer sheet). Subsequently, seven items were eliminated from the Elementary Student Questionnaire and two questions pertaining to student progress were added for field testing purposes. The items were: Have you learned a lot from this class? and Do you like to learn in this class? During this year, seven questions were also omitted from the Secondary Student Questionnaire and two new items on student progress were field tested: I have learned about the subject being taught in this class and I have enjoyed learning about the subject being taught in this class.

The Student Questionnaires remained unchanged for the 1986-87 year. The 4/10/87 minutes of the SCC indicate that in 1987-88 the K-2 questionnaire was eliminated because of administration difficulties and because it yielded only a small amount of data. Additionally, it did not accurately differentiate among levels of teachers, although it did not have a negative impact on candidates' scores (see 5/8/87 SCC minutes). Wording of the Elementary and Secondary Questionnaires was refined but no substantive changes were made. In 1988-89, the previously field-tested items were incorporated into the Elementary and Secondary instruments.

Competencies and Indicators. The Competencies and Indicators that serve as the bases for the items on the Student Questionnaires are the same as those underlying the other instruments. Hence the development procedures described for the Principal Questionnaire earlier in this chapter apply here also.

Administration of the Student Questionnaires. The Questionnaires are to be administered by the State Evaluator during the second visit. The teacher should not be present in the room where the Questionnaires are being administered. This practice helps reduce the potential for teacher influence on students' responses.

If the candidate's school is departmentalized, the Questionnaires are to be administered to a section or class of the candidate's choice, providing that a minimum of 20 students are included. In self-contained teaching situations, the entire class receives the questionnaire. Evaluators are instructed that it may take longer to administer the questionnaire to the lower grades but in no case should it take more than half-an-hour.

Procedures for administering the Student Questionnaires are similar for both forms. However, the elementary version requires that the evaluator read each item aloud while students respond. The evaluators are also instructed to check periodically for understanding. The manner in which they do so may or may not be uniform as described below.

Instructions in the Evaluator Training Manual (e.g., Tennessee State Department of Education, 1988-89) indicate that the evaluator should ensure that the teacher has provided each student with a #2 pencil, ask the teacher to leave the classroom, identify him/herself and the purpose for being in the class, give each student one copy of the Questionnaire, and count and record the number of questionnaires distributed.

The evaluator is additionally instructed to inform students that their participation is voluntary and that their identities will be kept confidential. If students are unwilling to participate, the evaluator is to retrieve the Questionnaire and ask the student to work quietly on something else while the Questionnaire is being completed by the other students. If students have questions about any items, the evaluator is to provide alternate phrasing or additional information. How this is done is left to the discretion of the evaluator and, hence, may not be uniform across evaluators. Collection procedures for completed Questionnaires are also up to the evaluator to decide so long as the number collected equals the number distributed.

Guidelines, in script form, for presenting Questionnaires to the students are provided in the Evaluator's Training Manual (e.g., Tennessee State Department of Education, 1988-89); but, the Manual also emphasizes that the guidelines are just that—guidelines—and that the the evaluators may reword, reorder, or even redo the guidelines to suit a particular group of students so long as the following elements are retained:

- Who is asking the students to complete the questionnaire (State Department of Education) and why (to help in the evaluation of their teacher).
- How to mark the sheet (#2 pencil, fill the bubbles).
- That this is voluntary and anonymous.
- Questions must be read aloud to students in grades 6 and below.

- Check for understanding at least once ("Any questions?").

Training/Implementation

As with the Principal Questionnaire, this section of the chapter is reliant upon extant information available in the files outlined in Appendix C. Training in Questionnaire-related procedures were not directly observed by members of the validation team. According to a technical paper presented by Malo (1987), training in Questionnaire administration is approximately four hours total including the Principal Questionnaire.

A review of training materials (e.g., Evaluator's Training Manual) (e.g., Tennessee State Department of Education, 1988-89) indicates that most of the instruction centers on the logistics of Questionnaire administration. In the case of the Student Questionnaires, procedures leading up to actual Questionnaire administration (e.g., asking teachers to leave the classroom, etc.) are straightforward; however, actual administration allows for evaluator discretion. For example, evaluators may even reword information to meet the perceived needs of particular students or classes. The extent to which Evaluators modify the guidelines contained in these training manuals is not assessable at this time.

Summary

The Questionnaires add to the credibility of the CLS process by fulfilling the important political need for input by stakeholders other than the candidates or State evaluators. By providing alternate data sources for the CLS Indicators, they additionally contribute to the potential validity of the CLS process and resultant decisions.

The instruments are reliable as demonstrated by their reliability coefficients. They undoubtedly suffer from the problems associated with using questionnaires as a data source for any purpose. That is, they are based on reports by individuals having particular perspectives rather than on relatively objective observations. Also, the vested interests of the respondents (e.g., principals whose ratings will ultimately be known to the candidates via the final evaluation report and who must continue to work with candidates regardless of the CLS decision) may contribute to reliable, but not particularly valid, ratings.

The extent to which administration of the Student Questionnaires allow for individual student differences (e.g., rewording for students who do not seem to understand items) is commendable; however, it also contributes to nonstandard administration. To the extent that administration is not consistent, the validity of responses may be questioned.

Chapter VI: Statistical Indicators of Validity for the Tennessee Career Ladder Program

Reports by French, Malo, and Rakow (1987) and by Rakow and Malo (1989) provide the results of fairly extensive analyses of several issues that are relevant to this report. In particular, those reports provide information on the internal consistency reliabilities of sub-scores that contribute to the composite Career Ladder score, interrelationships among various subscores and between subscores and the composite, and information on the degree to which the various subscores differentiate among Career Ladder Levels. In addition, they provide information on the weighting of various types of scores to form Domain scores and the weighting of Domain scores to form the composite scores that are used to determine qualification for Levels II and III on the Career Ladder.

We have reviewed those reports and make fairly extensive reference to them in our discussion of statistical evidence of validity of the Career Ladder System (CLS). We also have conducted our own data analyses using a data file that contained records for program participants in 1986, 1987, 1988, and 1989. These included correlational analyses involving domain and composite scores and analyses of the degree to which the domain scores differentiate the three Career Ladder Levels. The latter analyses provide an independent assessment of some of the issues addressed in the French, Malo, and Rakow and the Malo and Rakow reports. In addition, however, analyses were conducted to evaluate the sensitivity of the composite score to alternative weightings of the five Domains and to investigate the relationship of background characteristics, particularly race and gender, to Career Ladder Level attainment. Finally, the performance of recipients of various teacher awards (e.g., Teacher Education Association Distinguished Classroom Teachers) in the CLS was investigated.

Types of Scores and Determination of Career Ladder Level

In addition to meeting experience requirements and passing the communication test requirements, a minimum score of 600 is required on the Career Ladder composite to qualify for Level II and a minimum of 700 is required for Level III. To qualify for one of these levels, candidates are also required to have scores in each of the five Domains that are combined to form the Career Ladder composite of at least 450. The five Domains for which scores are reported are (I) Planning, (II) Teaching Strategies, (III) Evaluation, (IV) Classroom Management, and (V) Professional Development and Leadership. The scores for each Domain are reported on a 200 to 800 scale. To compute the composite, the five Domain scores are multiplied by weights of .15, .35, .15, .25, and .10 for Domains I through V, respectively, and summed.

With the exception of Domain V (Leadership), which does not include a consensus score, each Domain score is obtained as a weighted combination of a scaled data score (DATA), the principal score (PRIN), and the consensus score (CONS) for that domain. The weighted sums are defined as follows:

Domain I = .65(DATA 1) + .10(PRIN 1) + .25(CONS 1);
Domain II = .71(DATA 2) + .04(PRIN 2) + .25(CONS 2);
Domain III = .65(DATA 3) + .10(PRIN 3) + .25(CONS 3);
Domain IV = .63(DATA 4) + .12(PRIN 4) + .25(CONS 4);
Domain V = .80(DATA 5) + .20(PRIN 5).

The DATA scores for each Domain are based on equipercentile conversions of weighted sums of standardized scores from the relevant data sources for that domain. For Domains II, III, and IV the data sources consist of scores obtained from the Teacher Observations (OBS), the Dialogues (DIAL), the Student Questionnaires (STU), and the Professional Skills Test (PST). Domain I does not involve a Student Questionnaire so the DATA 1 score is based on a weighted combination of OBS, DIAL, and PST. The DATA 5 score is based on the only data source involved in that Domain, i.e., the Professional Development and Leadership Summary (PDLS).

Reliability

Rakow and Malo (1989) estimated internal consistency reliabilities using a modification of Cronbach's coefficient alpha. Estimates were not obtained for the PST or in cases where alpha is not applicable (e.g., the consensus scores or the dialogue score in the Classroom Management domain). The coefficients that were reported by Rakow and Malo are summarized in Table 1. With the exception of the DATA 3 score and the Domain V score, all of the internal consistency coefficients are .70 or higher, and most are in the .80 to .95 range.

High internal consistency is not a requirement for effective measures, especially when the interest is in domains that are relatively heterogeneous such as is the case in the Career Ladder Program. The magnitudes of the coefficients obtained are at reasonable levels and demonstrate that the subscores (e.g., OBS, DIAL) have relatively high internal consistency. Since the Domain scores, by design, combine several sources of information it is not surprising that internal consistences of the Domain scores are lower.

Intercorrelations of Subscores and Composite Scores

Rakow and Malo (1989) report a complete multitrait-multimethod intercorrelation matrix that includes correlations among the various subscores (e.g., OBS 1, DIAL 3, etc.) as well as the DATA, Domain, composite, and the actual Career Level attained. These intercorrelations provide a wealth of information for understanding the measurement characteristics of the CLS and we will consider some of these correlations in detail.

As was noted above, there are three or four subscores that are combined to form the DATA scores for Domains I through IV. It is instructive to consider the multitrait-multimethod correlation matrix shown in Table 2 for the four types of data subscores for Domains I through IV.

The idealized multitrait-multimethod matrix, when the variables are arranged as they are in Table 2, would have relatively high correlations in each of the four triangular segments of the table. These "validity triangles" contain the intercorrelations among the scores within a single Domain, (e.g. Planning) obtained by different methods. High correlations in these triangles provide evidence of "convergent validity," that is the degree to which different methods provide convergent evidence regarding the performance of a teacher within a domain. It can be seen, for example, that the Observation and Dialogue scores have moderately high correlations (.38 to .49) with each other in Domains I, II, and IV, and a somewhat lower correlation (.25) in the Evaluation Domain. Student Questionnaire scores are positively related, albeit at a generally lower level with the Observation and Dialogue scores. Scores on the Professional Skills Tests, with the possible exception of the Planning Domain, have little relationship to scores obtained from the other three domains.

The correlations in Table 2 that are shown in bold face type and underlined are based on scores for different domains obtained by the same method (e.g., Observation). If it is the domain rather than the method that is most important in determining an individual's scores, then those underlined correlations should generally be lower than the correlations in the validity triangles. Such a finding would provide evidence that the scores have "discriminant validity," that is, that the differential performance in the different domains can be distinguished by a given method. From an inspection of the correlations in Table 2, it is obvious that just the opposite usually holds for these data. That is, the highest correlations are obtained from the use of a common method across different Domains rather than the use of different methods within a single Domain. Scores obtained from Student Questionnaires, in particular, show essentially no differential validity for the different Domains since the across-domain STU correlations range from .83 to .89 while the STU scores correlate .32 or less with the scores obtained on a common domain using a different method.

Although the pattern of correlations in Table 2 provides only weak evidence of convergent validity and essentially no evidence of differential validity, such an outcome is not unusual. Method has often been found in other situations to dominate the magnitude of the intercorrelations, particularly in cases where ratings or judgments of observers are used. Students, in particular, seem to be providing data that reflect a single general characteristic rather than making sharp distinctions between strategies, evaluation, and classroom management.

The lack of discriminant validity would be a major concern if a single method were to be used to determine a teacher's level on the Career Ladder, or if decisions about teacher status were based on Domain scores rather than on the composite. Since neither of these conditions holds, we do not believe that the lack of discriminant validity is a serious problem for the CLS. It does suggest, however, that, for the future, consideration might be given to simplified methods of combining data (i.e., moving from the large number of original scores for the various components to a single composite score), and together with other findings reported below, that sharp distinctions or decisions based on domain score differences should be avoided.

In practice the OBS, DIAL, STU, and PST scores are not used separately for any decisions. Rather, these four data based scores are combined to form DATA scores, which are then used in combination with consensus and principal scores to derive DOMAIN scores. It is relevant to understanding the DATA scores, however, to consider the degree of relationship of those scores to the subscores from which they are derived.

Table 3 lists the correlations of the OBS, DIAL, STU, and PST subscores with the DATA scores for each Domain. These correlations provide an indication of the "effective" weights of the four subscores in determining the DATA score within each of these Domains. As can be seen, the PST is least related, and therefore has the smallest effective weight in determining the DATA scores in each of the four Domains. DIAL has the highest correlations with DATA for Domains I, II, and III, but both OBS and STU are more highly correlated with DATA than DIAL is for Domain IV. This suggests that the Dialogues are critical in deciding the Planning, Strategies, and Evaluation DATA scores, but not the Classroom Management DATA score. OBS is the most important data source for the latter DATA score and the second most important source for the other three.

As was previously indicated, once the DATA scores are computed they are combined with the principal (PRIN) and consensus (CONS) scores to obtain the Domain scores that are reported and, in turn, combined to form the final composite score. The weights used to compute the Domain scores which were reported earlier, are sometimes referred to as the "nominal" weights to distinguish them from the "effective" weights. The nominal weights are used to perform the calculations, but the importance of each subscore in determining the relative standing of candidates on the Domain score may or not be closely parallel to the nominal weight because subscore variances and intercorrelations also influence the effective weights. Thus it is useful to consider the correlations of the DATA, PRIN, and CONS subscores with the Domain scores to get a better idea of the effective weights of the subscores.

Correlations of the DATA, PRIN, and CONS subscores with the corresponding Domain score are listed in Table 4. The striking thing about the results in Table 4 is the extraordinarily high correlations between DATA and the overall Domain scores for all five Domains . The correlations of .97, .98, and .99 between DATA and DOMAIN for all of the Domains suggest that, for purposes of determining Domain scores, the principal and consensus scores are unnecessary. The principal subscores have relatively low correlations with the Domain scores in all five cases, especially when it is considered that these correlations reflect part-whole relationships.

As was stated above, the nominal weights for combining the Domain scores to obtain the overall composite scores are .15, .35, .15, .25, and .10 for Domains I through V, respectively. Once again, to evaluate the effective weights it is useful to consider the part-whole correlations of the domain scores with the composite scores. Those correlations were computed separately for the 1986, 1987, 1988, and 1989 data sets provided by the Tennessee Department of Education. Those correlations and the number of cases on which they are based are shown in Table 5. Also listed in Table 5 are the correlations reported by Rakow and Malo (1989) for Fall 1987 - Spring 1988 applicants included in their analyses. The latter data set presumably overlaps substantially with the 1988 data set we analyzed, but involved a somewhat smaller group of applicants.

The effective weights suggested by the correlations are generally consistent with the nominal weights. For every data set the highest correlation between a Domain and the composite score was obtained for the Strategies Domain which has the largest nominal weight (.35). Classroom Management consistently had the second highest correlation and it is the domain with the second highest nominal weight (.25). The lowest correlation is obtained consistently by the Domain with the lowest nominal weight, Professional Development and Leadership. It should be noted, however, that the Domain-composite correlation for that Domain has declined over time and was more discrepant, especially in 1987 and 1988, from the corresponding correlations for Domains I and III than would be expected from the nominal weights of .10 and .15.

A reason for the apparently low correlation between Domain V and the composite is that the Domain V scores have relatively low correlations with the other four Domain scores (see Rakow & Malo, 1989). Domain scores I through IV have relatively high intercorrelations, especially when compared to the internal consistency coefficients. Indeed, those Domain intercorrelations are high enough to suggest that differences in relative standing on the first four domains should be interpreted with caution.

Domain Score Contributions to Differentiation Among Career Ladder Levels

French, Malo, and Rakow (1987) and Rakow and Malo (1989) investigated the differentiation among Career Ladder Levels I, II, and III by computing standardized means of subscores for participants who achieved each level. The differences between means in standard deviation units provided an index of the degree to which each subscore contributed to the differentiation between Levels attained.

We conducted parallel analyses for Domain scores separately for participants in 1986, 1987, 1988, and 1989. Our results[4] are consistent with those reported earlier by French, Malo, and Rakow (1987) and by Rakow and Malo (1989). They show that all five Domain scores contribute to the differentiation among Levels attained. As would be expected, the sharpest differentiation among Levels is provided by Strategies, the Domain which receives the largest weight in the composite. Classroom Management, which has the second largest weight, does the second best job of distinguishing among Levels, and, as would be expected from its low weight and modest correlation with the composite, the Leadership Domain does not distinguish among the Levels as sharply as the other Domain scores do.

Sensitivity Analyses

By eliminating one Domain score at a time and rescaling the composite based on the weighted combination of the remaining four scores it is possible to get some sense of the degree to which CLS results are sensitive to the inclusion or exclusion of any particular Domain. The estimates of the alternate composite scores and levels for each combination of four of the five Domain scores are illustrated in Table 6.

As can be seen in Table 6, the illustrative candidate had a score just below the Level II minimum score, but was well above the minimum on the Domains that receive small weights (Domains III and V). Scores on both of the highly weighted Domains (II and IV), however, are below the 600 cutoff. Therefore, it is not surprising that the estimated scores based on the exclusion of one of the latter Domains exceed 600. It might also be noted that simply weighting all five Domains equally would have results in a score of 612 for this illustrative candidate.

Calculations such as those illustrated in Table 6 for a single candidate were performed for all candidates in the 1986, 1987, 1988, and 1989 data sets. The resulting alternative composites, labeled C1234, C1235, C1245, C1345, and C2345, where the numbers indicate the Domains included, were intercorrelated and used to investigate the relationship of background characteristics to alternate composites. The latter results will be considered later. Here we focus on the correlational results.

For the 1986 and 1987 data sets the correlation between the operational composite and any of the alternative composites was in every case rounded to .99 or 1.00. This suggests that changing to one of the alternative measures would yield essentially equivalent results. Correlations for the 1988 and 1989 data sets were lower, but still very high. The lowest correlation in both 1988 and 1989 was .96, a value that was obtained between C1345 and the operational composite. C1345 is the alternate composite obtained by eliminating Domain II, Strategies. Since Strategies was previously shown to have the largest nominal and effective weights, it is hardly surprising that elimination of that Domain score would yield an alternate composite that had the lowest correlation with the operational composite. The fact that even that radical change would still yield an alternate composite that is correlated at least .96 with the composite actually in use suggests that the results are not sensitive to the particular configuration of Domain scores and weights.

Background Characteristics Related to Career Ladder Levels

The 1986, 1987, 1988, and 1989 data files were used to investigate the relationship of background characteristics to CLS performance. Those files contained data designating racial-ethnic group status, gender, and division of the state for most participants. These data were used to investigate the possibility that the Career Ladder Level attained is related to the race, gender, or location in the state of the applicants.

Table 7 presents the number of candidates applying for either Level II or Level III status and the percent of those applicants who attained one of those Levels separately for Black and White applicants for each year and for the sum of the four years. Also shown is the number applying for Level III and the percent of candidates who attained Level III separately for Blacks and Whites. Since results for candidates that did not specify a racial-ethnic group or specified a group other than Black or White are not included, the number of candidates within any year is less than was available for the correlational analyses reported above.

As can be seen in Table 7, the success rate for Black applicants was less than that achieved by White applicants in every year. The discrepancies are larger for Level III applicants than for Level II applicants. There are many possible reasons for the differences in success rates for Black and White applicants, including group differences in educational preparation and in teaching experiences, that we were not able to investigate within the constraints of the present study. However, as is discussed below, we did investigate the degree to which the differences shown in Table 7 might be sensitive to the particular weighting of the Domain scores used to determine the composite score.

Table 8 reports results for gender that are parallel to those shown in Table 7 for Blacks and Whites. There it can be seen that female applicants, particularly for Level III, have a higher success rate than do their male counterparts.

Similar analyses were conducted by division of the State. Since the latter analyses did not reveal any sizable differences, those results are not reported here. District comparisons of the proportions of teachers at each Career Ladder Level were also made. With one exception, the proportion of Level III teachers did not vary greatly by district. Excluding the Northwest Tennessee District, which had 14.1 percent of its teachers at Level III in September, 1989, the percent of Level III teachers by district varied only by 1.8 percent, ranging from a low of 10.1 percent for the Upper Cumberland District to 11.9 percent for the East Tennessee District.

Some relatively large differences in the percentage of teachers at Career Ladder Levels II and III can be found, however, for counties that are at the extremes of the per capita income distribution. According to the Tennessee Statistical Abstract for 1989, only Williamson County, with a total of 488 teachers at Career Ladder Levels I, II, or III in September, 1989 had a per capita income greater than $17,000 in 1986, the most recent year reported in the Statistical Abstract. On the other hand, four counties (Fentress, Hancock, Lewis, and Pickett), with a total of 316 teachers at Career Ladder Levels I, II, or III, had per capita incomes below $7,000 in 1986. Roughly one teacher in four (27.5%) in Williamson County had attained either Level II or Level III status, compared to only about one teacher in ten (10.4%) across the four counties with the lowest per capita income. The corresponding figures for Level III are 15.1% for Williamson County and 7.0% for the combination of the four counties with the lowest per capita income.

Effect of Composite Definition on Group Differences in Passing Rates

In an attempt to investigate the possibility that the group differences in passing rates could be reduced by different definitions of the composite score, five alternate composites, each defined by excluding one of the Domain scores were used to estimate the failure rates by race and gender. Estimates of the percent of applicants for either Level II or Level III who would have attained the minimum standard of 600 if one of the alternative composites had been used are provided separately for Black and White applicants in Table 9.

It is evident from an inspection of Table 9 that the alternate composites would not markedly alter the magnitude of the group differences in passing rates. The estimated passing rates do differ somewhat for the different alternate composites, but the composites that have a higher passing rate for Black applicants generally also have a higher passing rate for White applicants.

Thus, there is no basis for suggesting one of these alternative composites for purposes of reducing the group differences in percent passing.

External Evidence of Validity

All of the results considered to this point have been based only on Career Ladder Indicators of teacher performance. External indicators of teacher excellence are difficult to obtain. One could imagine instituting another data collection system that would attempt to provide independent indicators of excellence of teaching against which the CLS could be validated. However, such a system would not only be expensive, but would be subject to the same questions about its validity that confront the CLS. Consequently, our best source of external evidence regarding the CLS comes from the performance of teachers who have been recognized as excellent by other programs when they participated in the CLS.

The performance of finalists in award programs and that of recipients of awards in the CLS was provided by the Tennessee Department of Education. Data for a total of 6 such programs involving anywhere from 7 to 129 Tennessee teachers were obtained. For each program, the percent of award recipients who participated in the CLS was computed. Among those who participated the percent that had a composite score exceeding the Level II minimum (600) and the percent with a composite score exceeding the Level III minimum (700) were calculated. The results of these calculations are summarized in Table 10.

As can be seen, slightly over two-thirds of teachers who were recognized by one of the six programs shown in Table 10 applied for Level II or Level III. This is a substantially higher rate of application than for the state as a whole, suggesting that teachers who are judged to be outstanding on grounds other than the CLS are also more likely to apply to the CLS than are their peers who have not been recognized.

More directly relevant to the validation of the CLS, however, is the performance of the recognized teachers on the CLS composite. Approximately 19 out of 20 of the award recipients who apply achieve scores that exceed the minimum for Level II and approximately 9 out of 10 achieve scores above the Level III minimum. When these numbers are contrasted with the passing rates shown, for example, in Tables 7 and 8 it is obvious that teachers who are judged outstanding by one of these programs are also much more likely to qualify for Career Ladder Levels II or III than are a cross section of teachers. The strong showing of award recipients in the CLS provides external evidence that the System measure is a valid indicator of outstanding teaching.

Concluding Comments

The measures used in the CLS have been demonstrated to have relatively high internal consistency. Although the Domains are not sharply distinguished by the various measures, this is not a serious shortcoming because decisions are based on the combined results from all five Domains. Sensitivity analyses suggest that CLS results are not very dependent on the particular combination of weights and measures. Thus, there is no basis for suggesting another combination of measures that would either improve validity or reduce the magnitude of group differences in attainment of Career Ladder Levels II or III.

There is a limited amount of statistical evidence on which to judge the validity of the CLS against external indicators of excellence in teaching. What evidence there is comes from the performance of recipients of awards in the CLS and that evidence is quite positive. The high participation rates and high passing rates of teachers have been recognized as outstanding on other grounds, together with the wealth of evidence supporting the content validity of the instruments that was discussed in earlier sections of this report, provide a strong basis of support for Tennessee's CLS.

Chapter VII: Conclusions—Summary of Findings, Related Issues and Recommendations

Findings and related issues regarding the content and statistical properties of specific instruments and procedures are detailed within the preceding chapters and summarized below. Additionally, this chapter considers those findings and issues which are more programmatic in nature and posits several recommendations. Most of our conclusions are directly related to the validity issues targeted by the study. However, we also felt obligated to include pertinent incidental findings. Together, the findings, issues, and recommendations speak to the validity of inferences made about teaching quality on the basis of composite Career Ladder System (CLS) data. Individual instruments and procedures contribute to, but cannot constitute, the sole bases for such inferences.

Summary

Development. Existing documentation justifies the conclusion that technically sound and professionally recognized procedures were used to identify the Domains and Competencies assessed by the CLS instruments. The evidence further suggests that the processes and methods used to modify and refine the various instruments and procedures over time have been appropriate, politically sensitive, and based on the results of data analyses, expert judgment, and feedback from CLS participants.

With few exceptions, substantive changes in the instruments or procedures have been relatively minor across time. This may be partly due to the fact that in the initial development stages, only those Competencies and Indicators rated favorably by at least 70% of surveyed teachers were retained. Also, subsequent revisions were based, in part, on candidate feedback.

The few, relatively major changes have included the 1985-86 replacement of the Candidate Interview and Portfolio with the Dialogues and PDL Summary, the replacement of the Superordinate Questionnaire with the Principal Questionnaire, and the elimination of the Peer Questionnaire. These changes were justified on the bases of data collection and analyses experiences, the judgment of nationally renowned experts, and CLS participant feedback. In short, the eliminated or replaced instruments lacked candidate credibility, inadequately contributed to differentiation among Career Ladder Levels, and/or were dysfunctionally cumbersome to administer or evaluate. The 1985-86 instrumentation changes added to the efficiency, credibility, and validity of the System without compromising the previously identified Competencies and Indicators.

In 1987-88, the K-2 Student Questionnaire was eliminated. Despite repeated modifications, it remained difficult to administer and continued to yield little meaningful data beyond that available from other, more reliable sources (e.g., grade 3-12 Student Questionnaires).

Content. Given appropriate training and implementation, results yielded by the instruments should contribute to valid inferences regarding teaching quality. The instruments' Competencies and Indicators are based on effective teaching research, consensus by Tennessee teachers, expert judgment, and the Commission's reasoned approval. Also, the content of the instruments is generally consistent with teacher evaluation systems developed over the last decade. This general consistency does *not* imply that Tennessee's needs could have been met by adapting another system.

Evaluator Training and Selection. There is strong evidence that the content of training in the Dialogue, Observation, and PDL instruments is appropriate and adequate. The training has been observed to be of sufficient duration and depth to provide a basis for reliable use of the instruments and related procedures. By comparison, the need for extensive training in Questionnaire administration is minimized by the straightforward instructions contained in the Evaluator Training Manual (e.g., Tennessee State Department of Education, 1988-89). Nonetheless, actual administration allows for discretion on the part of the evaluator--discretion which seems reasonable but which may also influence the reliability of results. The extent to which evaluators actually modify the administration guidelines contained in the Evaluator Training Manual is not assessable at this time.

The criteria and procedures for selecting evaluators are basically sound. The application criteria include teaching experience, subject matter and grade level expertise, writing ability, and communication skills. All of these seem both reasonable and necessary basic qualifications which, along with an adequate training program, should yield competent evaluators.

The current process used for selecting evaluators from among the pool of qualified applicants--a pool that calls for appropriate representation by race, sex, and geographic region--requires that (1) all applications be read by two Department of Education staff and rated according to the qualifying criteria and (2) highly rated candidates be personally interviewed by two members of the staff. This process is reasonable and responsive to the need to fairly interview large numbers of candidates within short and intense timeframes.

Reliability and Implementation. In order to become certified, trainee evaluators must meet the criteria for reliability. The criteria are based on 80% agreement with pre-established expert judgment. Hence, it may be concluded that certified evaluators are able to reliably apply the Dialogue, Observation, and PDLS instruments.

The extent of interrater reliability based on actual classroom observations, rather than on training tapes, is not known. Dual-observer classroom observations are no longer used, one reason being that teachers perceived two observers as a signal that the System was not reliable.

Also unknown is the extent to which a candidate's writing ability is confounded with the quality of the activities assessed by the PDLS. Given the virtual unavoidability of such confounds, several controls have been built into the System (e.g., independent readings and ratings by two evaluators who must then reach consensus).

The Questionnaires are reliable as demonstrated by their internal consistency coefficients. However, the validity of the Principal Questionnaire must be interpreted with caution since the principals' ratings are individual, identifiable from other data sources, and ultimately known to the candidate. Despite this caution, and despite the fact that the results of the Questionnaires add little to the overall judgment of teaching quality, their inclusion is politically expedient. Also, there is no evidence that such inclusion negatively impacts CLS decisions.

Programmatic Findings/Issues

Psychometric Properties. Insofar as applicable, internal consistency reliability has been established for each of the CLS Domains: Planning, Teaching Strategies, Evaluation, Classroom Management, and Professional Development and Leadership. Although high internal consistency is not a requirement for effective measures, particularly composite measures based on relatively heterogeneous subscores, Table 1 indicates that the magnitudes of the five Domain coefficients are at reasonable levels. The subscores on which the Domain scores are based have relatively high levels of internal consistency.

Multitrait-multimethod methodology provided weak evidence of convergent validity among Domain scores and no evidence of differential validity. Given the rater-reliance of the various scores making up the composite, these findings are not particularly unusual. Rather, they suggest the importance of using more than one evaluation method to arrive at a composite as well as the importance of basing decisions regarding CLS attainment levels on composite, rather than Domain, scores.

Weighting and Scoring. Correlations between the data scores within each Domain (data scores being nominally weighted composites based on scores yielded by portions of individual instruments) and their corresponding domain scores are extremely high (.97-.99). These correlations suggest that the Principal Questionnaire and consensus scores, which combine with the DATA scores to form Domain scores, are technically unnecessary. However, their inclusion adds to the credibility of the process and does not detract from its validity.

The Domain scores for all domains except Professional Development and Leadership have relatively high intercorrelations. These findings suggest that relative standings across Domains should be interpreted with caution.

All five Domain scores contribute to the differentiations among CLS Levels with the greatest contribution being made by the Teaching Strategies Domain which carries the largest nominal weight (.35) and the least contribution being made by the Professional Development and Leadership Domain which carries the smallest nominal weight (.10). Analyses of composite scores that were based on the systematic elimination of single Domain scores and weighted combinations of the remaining scores indicates that similar decisions regarding CLS attainment Levels would result. Hence, while all Domain scores positively contribute to the composite score, eliminating single Domain scores and redistributing the weight of the missing Domain score among the remaining Domains does not differentially affect CLS Level attainment decisions.

Group Differences in Passing Rates. The data indicate that Black teachers have been somewhat over-represented at Career Ladder Level I and under-represented at Levels II and III. These findings may reflect the national shortage of minority teachers, the age/experience distribution of Tennessee teachers according to race, or the high percentage of candidates who did not specify their race on the data forms. Regardless of the reasons, the success rate for Black applicants was less than that achieved by White applicants in every year (1986-1989).

There are also gender differences in passing rates. Female applicants, particularly for Level III, have a higher success rate than do their male counterparts.

There is relatively little variation in passing rates across the major geographic regions of the state. However, relatively large differences in the percentage of teachers at Career Ladder Levels II and III exist with respect to per capita income. In a county having an average per capita income of $17,000 in 1986, 27.5% of the teachers had attained either Level II or Level III status. In four counties having an average per capita income of less than $7,000, only 10.4 % of teachers had attained either Level II or Level III status. At Level III, the attainment rates were 15.1% and 7.0%, respectively.

Alternative composite scores, based on exclusion of one of the Domain scores, were used to estimate success rates by race and gender. As shown in Table 9, the alternative composites would not substantially change the degree of success.

External Criteria. Teachers judged as outstanding on grounds other than the CLS are more likely to apply to the CLS than their unrecognized peers. They are also more likely to achieve Level II or Level III status, indicating that the CLS is a valid indicator of teaching excellence.

Conclusions/Recommendations

- The comprehensiveness and professional delivery of the evaluator training component of the CLS is laudable. It is also too dependent upon individual trainers; hence, its transportability is limited. While the training manuals and materials are impressively extensive, the training procedures, themselves, require detailed documentation. This is particularly important given the recent loss of staff involved in the conceptualization and on-going development of the System. Critical to this documentation process is the need to explicitly specify the decision rules for the reliability criteria used for training and certification purposes.
- Evaluator self-reports indicate that their training and experience has enriched and enhanced their classroom performance. Their individual and collective knowledge

and experience warrants consideration as a potential resource for furthering Tennessee's educational goals.

- Inferences based on individual CLS Instruments or Domain scores are not valid for judgments about attainment Levels. Such decisions should be based on composite scores only.
- The performance differences with regard to race and sex (1) underscore the importance of the evidence provided for external and content validity, (2) may warrant additional explanatory investigation, and (3) suggest the need for active recruitment and program marketing. At present, staff seem too lean and busy to market the true nature of the System.
- The content of the CLS has implications for the content of Tennessee's teacher education programs. Ways in which the State Department of Education and institutions of higher education can work together to recruit and provide excellent teachers merits serious consideration.
- Throughout the evaluator training sessions, on-site computer hardware and software should be available to increase the efficiency of scoring and timely feedback.
- Additional, high quality training tapes are needed to increase the validity with which the tapes may be used to make inferences about evaluator reliability as it relates to ensuing field experiences. Since the reliable rating of training tapes does not ensure reliable classroom

observations, etc., the re-institution of site-based reliability checks warrants consideration. Site-based reliability checks could provide stronger support that procedures are reliably implemented.

- At present, the post observation interviews actually require an evaluation of what was observed which is not the intent of the System. That is, the System's intent is that evaluations be based on multiple data points. Furthermore, these interim "mini" evaluations have the potential to misguide candidates by providing information based on a particular circumstance that will not necessarily generalize in beneficial ways to future observations.

Within the confines of the legislation, the enormity of the task, and the available resources, CLS staff have done a stellar job of conceptualizing, developing, and implementing a model teacher evaluation system that allows for valid inferences about teaching quality.

REFERENCES

AERA (American Educational Research Association), APA (American Psychological Association), & NCME (National Council on Measurement in Education). (1985). Standards for educational and psychological testing. Washington, DC: American Psychological Association.

Bird, T. (1990). The schoolteacher's portfolio: An essay on Possibilities. In: Millman & Darling-Hammond (Eds.), The new handbook of teacher Evaluation. Newbury Park, CA: Sage Publications.

Career Ladder Retreat. (1985, July 8). Status Report, Review and revision of the Teacher and Administrator Evaluation Systems.

Cornett, L. (1990). Paying for performance-Important questions and answers: The 1989 SREB Career Ladder Clearinghouse Report. Altanta, GA: Southern Regional Education Board.

French, R. (1983, October 20). Analysis of Teacher Review of Evaluation Competencies. Memorandum to members of the Interim Certification Commission. (File I-C-7-a).

French, R., Malo, G., & Rakow, E., (1987). Update on data analysis of Career Ladder Evaluation System for Teachers: Preliminary report. SCC Minutes Attachment II.A (5/8/87).

Furtwengler, C.C. (1987, April). Lessons from Tennessee's Career Ladder Program. Educational Leadership, 66-69.

Furtwengler, C.B. (1985). Evaluation procedures in the Tennessee Career Ladder Plan. Paper presented at the Annual Meeting of the American Education Research Association, Chicago, IL.

Glass, G. (1990). Using student test scores to evaluate teachers. In: Millman & Darling-Hammond (Eds.), The new handbook of teacher evaluation. Newbury Park, CA: Sage Publications.

Haertel, E. (1990). Setting standards on teacher certification tests. In: Millman & Darling-Hammond (Eds.), The new handbook of teacher evaluation. Newbury Park, California: Sage Publications, 1990.

Interim Certification Commission. (1984a, May 1-2). Minutes of the Interim Certification Commission.

Interim Certification Commission. (1984b, July 20). Minutes of the Interim Certification Commission.

Interim Certification Commission. (1985, March 15). Minutes of the Interim Certification Commission.

Interim Certification Commission. (1986, April 7). Minutes of the Interim Certification Commission.

Interim Commission. (1983, December 14-15). Career Teacher Evaluation System. (Files II-A-1-1 and I-C-8-e).

Kelly, J. (1989). Strengthening the teacher profession through national certification. In: New directions for teacher assessment. Proceedings of the 1988 ETS Invitational Conference. Princeton, NJ: Education Testing Service.

Malo, G.E. (1987, April). Training and reliability of peer evaluation: The Tennessee Career Ladder Program for Teachers. Symposium presentation, Annual Meeting of the American Educational Research Association, Washington, DC.

Malo, G., & French, R. (1987, February). The Tennessee Career Ladder: What it is and how it has changed. Thresholds in Education.

Millman, J. (1985, Summer). Recommendations regarding Career Ladder revisions. (File I-E-1-a-(8); Tennessee State Department of Education file).

National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform. Washington, DC: U.S. Department of Education.

Parkison, E. (1987, August). Changes in teaching behavior brought about by experience as evaluators in the Tennessee Career Ladder Program. Dissertation submitted to the Graduate School of Tennessee State University.

Rakow, E.A. & Malo, G. (1989, March). Technical measurement issues in teacher evaluation: The Tennessee experience. Paper presented at symposium for the National Council on Measurement in Education, San Francisco, CA.

Rosenholtz, S.J., & Smylie, M. A. (1983a, June). Master Teacher Program, Suggested Criteria for Master Teacher Selection and Evaluation. Report to the Commissioner, Tennessee State Department of Education. (File I-C-1-c).

Rosenholtz, S.J., & Smylie, M.A. (1983b, December). Effective teaching strategies: A synthesis of classroom research. Report to the Commissioner, Tennessee State Department of Education Master Teacher Program.

Shulman, L. (1989). The paradox of teacher assessment. In: New directions for teacher assessment. Proceedings of the 1988 ETS Invitational Conference. Princeton, NJ: Education Testing Service.

State Certification Commission. (1982, 19 June). Minutes of the State Certification Commission.

State Certification Commission. (1986, August 19). Minutes of the State Certification Commission. (File I-E-1-b(2)).

State Certification Commission. (1987, June 19). Minutes of the State Certification Commission. (File I-E-1-c(6)).

State Certification Commission. (1987, May 8). Minutes of the State Certification Commission.

Stodolsky, S. (1990). Classroom observation. In: Millman & Darling-Hammond (Eds.), The new handbook of teacher evaluation, Newbury Park, CA: Sage Publications.

Tennessee State Department of Education. (1988-89). Career Ladder technical manual: A guide to interpreting your scores. Teacher edition. Nashville, TN: Author.

Tennessee State Department of Education. (1988-89). Evaluator training manual. Nashville, TN: Author.

Tennessee State Department of Education. (1986) Planning grant to develop teacher incentive stimulators. Submitted to the U.S. Department of Education. Nashville, TN: Author.

Tennessee State Department of Education. (1984-90). Career Ladder teacher orientation manuals. (1984-85 through 1989-90) Nashville, TN: Author.

Tennessee State Department of Education, Career Ladder Division. (1986, September). Planning Grant to Develop Teacher Incentive Structures. Final Report. (File I-C-10).

Timpane, M. (1989). Will improving teacher assessment improve the education of children? In: New directions for teacher assessment. Proceedings of the 1988 ETS Invitational Conference, Princeton, NJ: Education Testing Service.

Urbanski, A. (1989). New directions for the career of teaching–the Rochester Experiment. In: New directions for teacher assessment. Proceedings of the 1988 ETS Invitational Conference. Princeton, NJ: Education Testing Service.

Table 1

Weighted Coefficient Alpha Reliability Estimates

(Based on Rakow and Malo, 1989, Table 3)

	Domain				
Source	(I) Planning	(II) Strategies	(III) Evaluation	(IV) Management	(V) Leadership
OBS	.80	.83	NA	.75	NA
DIAL	.90	.92	.87	NA	NA
STU	NA	.88	NA	.84	NA
PDL	NA	NA	NA	NA	.89
DATA	.83	.86	.62	.79	.89
PRIN	.94	.95	.81	.88	.78
DOMAIN	.70	.85	.74	.77	.59

Table 2

Multitrait-Multimethod Correlation Matrix for Data Subscores in Domains I Through IV (Decimal Points Omitted)

(Based on Rakow and Malo, 1989, Table 3)

	(I) Planning				(II) Strategies				(III) Evaluation				(IV) Management			
	OB	DI	ST	PST	OB	DI	ST	PST	OB	DI	ST	PST	OB	DI	ST	PST
OB1	--															
DI1	49	--														
ST1	NA	NA	--													
PST1	21	20	NA	--												
OB2	**68**	43	NA	09	--											
DI2	52	**34**	NA	20	43	--										
ST2	28	15	NA	04	26	19	--									
PST2	19	18	NA	**46**	04	15	01	--								
OB3	**44**	22	NA	04	**58**	23	13	03	--							
DI3	52	**30**	NA	18	39	**34**	22	18	25	--						
ST3	24	14	NA	10	25	18	**83**	02	13	25	--					
PST3	11	13	NA	**48**	00	12	-03	**42**	-02	16	04	--				
OB4	**62**	39	NA	14	**60**	44	29	06	**36**	41	25	01	--			
DI4	45	**25**	NA	13	34	**29**	25	12	25	**74**	24	07	38	--		
ST4	29	15	NA	06	29	21	**89**	02	20	25	**84**	-01	32	27	--	
PST4	16	19	NA	**47**	07	16	02	**43**	04	19	04	**41**	10	14	03	--

Table 3

Correlations of Observation, Dialogue, Student, and Skills Test Subscores with the Combined DATA Score Within Domains I Through IV (Based on Rakow and Malo, 1989, Table 3)

	DATA 1 Planning	DATA 2 Strategies	DATA 3 Evaluation	DATA 4 Management
OBS	.72	.71	.58	.74
DIAL	.86	.77	.74	.41
STU	NA	.44	.46	.70
PST	.39	.35	.30	.32

Table 4

Correlations of Data, Principal, and Consensus Subscores
with the Domain Score within Each Domain
(Based on Rakow and Malo, 1989, Table 3)

Subscore	DOMAIN I Planning	DOMAIN II Strategies	DOMAIN III Evaluation	DOMAIN IV Management	DOMAIN V Leadership
DATA	.99	.99	.98	.97	.99
PRIN	.23	.24	.24	.31	.29
CONS	.70	.78	.75	.79	NA

Table 5

Correlations of Domain Scores with the Composite Scores in 1986, 1987, 1988, and 1989

Year	N	Domain I Planning	Domain II Strategies	Domain III Evaluation	Domain IV Management	Domain V Leadership
1986	1743	.88	.96	.91	.94	.78
1987	1765	.84	.95	.88	.91	.69
1988	1201	.79	.93	.79	.84	.46
1989	845	.78	.91	.85	.87	.47
Rakow*	1032	.78	.93	.80	.84	.44

* Based on Rakow and Malo, 1989, Table 3. There data were fall 1987 and spring 1988 and are presumably a subset of data for 1201 analyzed here with 1988 data.

Table 6

Illustration of Estimating Career Ladder Level Based on Four of Five Domain Scores

Domain	Score	Weight	Weighted Domain Score
I	520.50	.15	78.08
II	578.88	.35	202.61
III	686.25	.15	102.94
IV	548.25	.25	137.10
V	724.80	.10	72.48
Total Score			593
Career Ladder Level*			I

Estimated Total Scores Using Each Combination of 4 of 5 Domain Scores

Domains Included	Domain Excluded	Estimated Score**	Level
I, II, III,IV	V	579	I
I, II, III,V	IV	608	II
I, II, IV, V	III	577	I
I, III,IV, V	II	601	II
II,III,VI, V	I	606	II

* Minimum score for Level II is 600 and minimum for Level III is 700.

** Estimated scores are calculated by summing the indicated weighted domain scores and dividing by the sum of the weights of the corresponding domains.

Table 7

Number of Applicants for Career Ladder Level III, for Levels II or III, and Percent of Applicants Attaining Those Levels by Year and Race

	1986	1987	1988	1989	Total
	Black Applicants				
Number Level III	140	146	84	56	426
Percent Attained	44.3	24.0	20.2	23.2	29.8
	White Applicants				
Number Level III	964	968	600	466	2998
Percent Attained	59.0	42.5	41.3	40.1	47.2
	Black Applicants				
Number Level II or III	197	204	116	73	590
Percent Attained	75.6	61.8	65.5	72.6	68.5
	White Applicants				
Number Level II or III	1456	1447	989	705	4597
Percent Attained	87.9	84.2	88.4	86.5	86.6

Table 8

Number of Applicants for Career Ladder Level III, for Levels II or III, and Percent of Applicants Attaining Those Levels by Year and Gender

	Male Applicants				
	1986	1987	1988	1989	Total
Number Level III	300	289	160	119	868
Percent Attained	47.7	27.0	24.4	29.4	34.0
	Female Applicants				
Number Level III	843	879	568	433	2723
Percent Attained	60.3	43.2	42.3	38.9	47.6
	Male Applicants				
Number Level II or III	414	431	252	165	1262
Percent Attained	79.2	71.7	77.0	77.6	76.0
	Female Applicants				
Number Level II or III	1294	1313	929	663	4199
Percent Attained	88.5	83.1	87.4	87.0	86.3

Table 9

Estimated Percent of Applicants for Career Ladder Levels II or III Who Would Have Attained the Level Applied for if Alternative Composites Were Used

Alternative Composite	1986	1987	1988	1989
Black Applicants				
C1234	79.7	61.8	63.8	69.9
C1235	78.7	60.8	66.7	71.2
C1245	80.7	67.6	66.4	71.2
C1345	81.8	69.6	74.1	78.1
C2345	81.7	72.1	75.0	78.1
Current Composite	75.6	61.8	65.5	72.6
White Applicants				
C1234	86.5	82.4	86.2	83.8
C1235	87.3	81.3	84.8	84.7
C1245	87.4	84.6	88.1	85.7
C1345	88.6	85.6	90.9	87.2
C2345	88.7	86.6	90.6	88.9
Current Composite	87.9	84.2	88.4	86.5

Table 10

Career Ladder Application and Attainment for Tennessee Teachers Who Have Received Other Teaching Awards

Award	Number of Recipients	Percent Recipients Applying	Percent Applicants at Level II or III	Percent Applicants at Level III
Tenn. Teacher of Year Dist. Finalist	129	67.4	97.7	93.1
TEA Distinguished Teachers	53	56.6	100.0	93.3
Tenn. Nat. Candidates Presidential Award	24	87.5	100.0	90.5
Tenn. For. Language Assoc. Outstanding Teachers of Year	7	71.4	100.0	100.0
Health, Phys. Ed., Rec. & Dance Assoc. Awards	10	90.0	44.4	44.4
Tenn. NASA in Space Semi-Finalists	7	71.4	100.0	100.0
Total Awards Above	230	68.3	95.5	90.4

Appendix A: Tennessee Code Annotated, Section 49-5-5103(4)
# of the Comprehensive Education Reform Act

49-5-5103. Duties of commission – Local evaluation plans. – The state certification commission shall have the following duties:

(4) To recommend, in consultation with the advisory commission on teacher education and certification, to the state board of education certification and evaluation standards, criteria and procedures, including education and competency requirements, for use by the state certification commission, the regional certification commissions and any local education agency administering its own evaluation procedures, criteria and instruments which have been approved by the state board of education under the provisions of parts 50 through 57 of this chapter. Following the state board of education's approval of these standards, criteria and procedures, they, and any necessary rules shall be promulgated in accordance with the provisions of pats 50 through 57 of this chapter and the Uniform Administrative Procedures Act, title 4, chapter 5. Copies of these standards and criteria shall also be filed with the standing education committees of the senate and house of representatives. The policies, standards and rules regarding evaluation standards, procedures, criteria and instruments shall be the responsibility of the state board of education acting upon the recommendation of the state certification commission and shall not be subject to alteration or limitation by whatever means. The criteria for the evaluation of educators shall be validated and tested to eliminate racial or sexual bias prior to its use by the state certification

commission and the regional commission. It is the intent of the general assembly and the requirement of the career ladder programs that the procedure of evaluation assure the educator a fair, unbiased and objective determination of professional competency and that no procedure of evaluation be adopted and no certification or other decision hereunder be made or withheld which may discriminate or exclude an educator on the basis of race or sex, and that such procedure, including but not limited to such criteria specifically mention and be directed toward prevention of such discrimination or exclusion on account of race or sex. The state certification commission shall report to the state board of education and the legislative oversight committee annually on the validation and testing of evaluation criteria, including names of consultants, procedures, instruments, and results used to assure that educators receive a fair, unbiased and objective determination of professional competency.

Appendix B: National Educational Goals

National Educational Goals

By the year 2000, all children in America will start school ready to learn.

Families, Communities and Children's Learning;
Cultural Diversity and Second Language Learning;
Assessment, Testing, and Evaluation;
Research on Dissemination and Knowledge Utilization

By the year 2000, the high school graduation rate will increase to at least 90 percent.

Education in the Inner Cities;
Organization and Restructuring of Schools;
Education Policies and Student Learning;
Assessment, Testing, and Evaluation;
Research on Dissemination and Knowledge Utilization

By the year 2000, American students will leave grades four, eight, and twelve having demonstrated competency in challenging subject matter including English, mathematics, science, history, and geography; and every school in American will ensure that all students learn to use their minds well, so they may be prepared for responsible citizenship, further learning, and productive employment in our modern economy.

Student Learning;
Writing and Literacy;
Postsecondary Learning, Teaching and Assessment;
Literature Teaching and Learning;
Assessment, Testing, and Evaluation;
Research on Dissemination and Knowledge Utilization

By the year 2000, U.S. Students will be first in the world in science and mathematics achievement.

Mathematics Teaching and Learning;
Science Teaching and Learning;
Assessment, Testing, and Evaluation;
Research on Dissemination and Knowledge Utilization

By the year 2000, every adult American will be literate and will possess the knowledge and skills necessary to compete in a global economy and exercise the rights and responsibilities of citizenship.

Adult Literacy;
Educational Quality of the Workforce;
Assessment, testing, and Evaluation;
Research on Dissemination and Knowledge Utilization

By the year 2000, every school in America will be free of drugs and violence and will offer a disciplined environment conducive to learning.

Education in the Inner Cities;
Cultural Diversity and Second Language Learning;
Learning to Teach;
Families, Communities, and Children's Learning;
Education Finance and Productivity;
Teacher Performance Evaluation and Educational Accountability;
Assessment, Testing and Evaluation;
Research on Dissemination and Knowledge Utilization

Appendix C: Outline of Supporting Material Available for Proposed Activities to Validate Career Ladder Evaluation System

OUTLINE OF SUPPORTING MATERIAL AVAILABLE FOR PROPOSED ACTIVITIES TO VALIDATE CAREER LADDER EVALUATION SYSTEM

I. Validity in Content and Process

A. Purpose of the program

1. Mandated purpose of the Career Ladder Program

   a. Comprehensive Education Reform Act of 1984

   ...to establish a new professional career ladder program for full time teachers, principals and supervisors [TCA 49-5-5002].

   ...The new career ladder program shall be designed to promote staff development among teachers, and to reward with substantial pay supplements those teachers evaluated as outstanding and who may accept additional responsibilities are applicable [TCA 49-5-5002(b)(1)].

   ...The new career ladder shall be designed to improve the administrative skills of principals, and reward with substantial pay supplements those principals evaluated as outstanding [TCA 49-5-5002(b)(2)].

   ...The new career ladder shall be designed to improve the administrative skills of assistant principals and to reward with substantial pay supplements those assistant principals evaluated as outstanding [TCA 49-5-5002(b)(3)].

   ...The new career ladder shall be designed to improve the skills of administrative supervisors, and reward with substantial pay supplements those supervisors evaluated as outstanding [TCA 49-5-5002(b)(4)].

   Further:

   (e) It is the intent of the general assembly that the salary supplements provided for herein be awarded on the basis of outstanding performance and that that standards utilized for this purpose be kept high. The commissioner of education shall report to the committee created in 3-15-301 on the adequacy of the standards [TCA 49-5-5002(e)].

   b. Report of the Select Committee on Education, Tennessee General Assembly, January 1984

2. Other stated purposes

a. Speeches by Governor Lamar Alexander

...an incentive pay system that will make teaching a fully professional career, draw our best young people into it, keep our best teachers in it, challenge our best teachers to do even better, and inspire excellence in our classrooms by rewarding excellence in our teachers.

Goal: To improve the quality of elementary and secondary education in Tennessee by strengthening the knowledge, preparation, incentives and rewards of classroom teachers.

Specific Objectives:

(1) To foster professional growth and awareness among educators.

(2) To encourage educators to assume greater responsibility for educational quality.

(3) To facilitate the development and enforcement - by educators - of higher standards of performance for themselves, their professional colleagues and their students alike.

(4) To make the work of the teacher and the principal more rewarding and satisfying as professional careers.

(5) To improve the effectiveness of schools by strengthening their educational leadership.

[Tennessee's Better Schools Program, Governor Lamar Alexander's State of Education Address, January 28, 1983].

A pay system that will contribute toward the goal of making public school teaching truly a professional career.

A pay system that will attract our best and brightest young people into the educational field.

A pay system that will keep our best teachers in their chosen profession.

A pay system that will help achieve excellence among our students by rewarding excellence in our teachers [Homework on Tennessee's Master Teacher Pay Plan, Summary of Presentation to the President of the United States by Governor Lamar Alexander, Farragut High School, Knoxville, Tennessee, June 14, 1983]

...to establish a Career Teacher program that will help keep and attract the best public school teachers in America, a program that will give every single teacher in Tennessee a chance to earn a salary that is $10,000 higher than the best paid teacher in his or her school district makes today; [The Message of Governor Lamar Alexander presented to the Ninety-Third General Assembly of the State of Tennessee (Extraordinary Session), January 10, 1984]

3. Pertinent Data

a. Analysis of Differentiation Among Career Ladder Levels, Student Progress Items (SCC Minutes Attachment II.A., 6/10/88]

b. "The Commissioner's Report on Education," Draft, Submitted to State Board of Education, January, 1986

c. "Student Teacher and School Performance," Second Annual Report Submitted to the Governor and the General Assembly by the Board of Education and the Commissioner of Education, January 8, 1987

d. "Student Teacher and School Performance," Third Annual Report Submitted to the Governor and the General Assembly by the Board of Education and the Commissioner of Education, January 29, 1988

e. "Student Teacher and School Performance," Fourth Annual Report Submitted to the Governor and the General Assembly by the Board of Education and the Commissioner of Education, January 27, 1989

f. Tennessee Student Test Results 1987-1988, 1988-1989

Highlights of these reports indicate that achievement by Tennessee students has increased steadily since 1985 on the Stanford Achievement Test in Grades 2, 5, and 7. What effect the Career Ladder Program has had on this improvement would be extremely difficult to assess.

g. Doctoral Dissertations

h. Newspaper Clippings

i. A study of the Career Ladder and Extended Contract Programs, December 1988

j. Professional Articles

B. Expertise and training of consultants, panels, and sounding boards:

1. Consultants

a. State Department of Education Staff Activities for Interim Commission, November-December 1983

b. Dr. W. James Popham - Vita
Professor, College of Education
University of California at Los Angeles and
Director, Instructional Objectives Exchange
Los Angeles, California

c. Dr. Robert Soar - Vita
Professor, College of Education
University of Florida
Gainesville, Florida

d. Dr. Jason Millman - Vita
Professor, College of Education
Cornell University
Ithaca, New York

e. Dr. Susan J. Rosenholtz - Vita
Assistant Professor, Education and Sociology
Peabody College of Vanderbilt University
Nashville, Tennessee

f. Mr. Mark Smylie
Research Assistant
Vanderbilt Institute for Public Policy Studies
Nashville, Tennessee

g. Dr. Jane Stallings
Professor of Education
Peabody College of Vanderbilt University
Nashville, Tennessee

h. Dr. Lester Soloman
Associate Director for Performance-based Certification
Division of Staff Development
Georgia Department of Education
Atlanta, Georgia

i. Mr. Paul Hersey
Director, National Secondary School Principals'
Assessment Center
Reston, Virginia

j. Dr. George Redfern
Consultant, Personnel Education
Educational Research Services
Arlington, Virginia

k. Dr. Lawrence M. Aleamoni
Professor, Department of Educational Psychology
University of Arizona
Tucson, Arizona

l. Madeline Hunter-Team
Staff Development Training Program
Rowland Unified School District
Rowland Heights, California

m. Anderson, Niebuhr & Associates, Inc.
1885 University Avenue
St. Paul, Minnesota

n. Appalachia Educational Laboratory
P.O. Box 1348
Charleston, West Virginia

o. Dr. Carolyn Evertson
Peabody College of Vanderbilt University
Nashville, Tennessee

2. Study Councils and Statewide Organizations

a. Teachers' Study Council
[July - December, 1983]

b. Better Schools Task Forces [statewide]
[1983 - 1985, e.g., Memphis City Schools]

c. Deans of the Tennessee Colleges of Education
[February, 1983]

d. Higher Education Task Forces
[August - December, 1983]

e. Tennessee Association for Supervision and Curriculum Development
{March, 1983]

f. Principals, Supervisors, and other Central Office Administrators
[Nine district meetings were held across the state during May and June, 1983. Between 800 and 1,000 principals and supervisors attended and were asked to provide input.]

- g. Superintendents Study Council
  [Regional meetings were held in May, 1983 seeing input.]
- h. Supervisors' Study Council
- i. Principals' Study Council
- j. Tennessee Association of Secondary School Principals
- k. Tennessee Association of Elementary School Principals
- l. Tennessee Educational Foundations Association
- m. Tennessee Organization of School Superintendents [TOSS]
- n. Tennessee School Board Association [TSBA]
- o. Tennessee Education Association [TEA]
- p. State Parent Teachers' Association [PTA]
- q. Teacher Associations [content related]
- r. Summary - Better Schools Meetings 1983
- s. Better Schools Reports, February - April 1983

3. Panels, Task Forces, Sounding Boards, Committees

- a. Attendance Supervisors Task Force
  [Fall, 1983]
- b. Audiologists Task Force
  [Fall, 1983]
- c. Comprehensive Developmental Classes Task Force
  [Fall, 1983]
- d. Consulting Teacher/Diagnostician Task Force
  [Fall, 1983]
- e. Resource/Chapter I Task Force
  [Fall, 1983]
- f. School Counselors Task Force
  [Fall, 1983]
- g. School Psychologists Task Force
  [Fall, 1983]

h. School Social Workers Task Force
[Fall, 1983]

i. Speech and Language Specialist/Pathologist Task Force
[Fall, 1983]

j. Professional Skills Testing Committee
[January, 1984]

k. Administrator/Supervisor Scale Development Panel
[January 23-27, 1984]

l. Career Ladder Teacher Evaluation Instrument and Standards Review Panel
[August 1-2, 1984]

m. Career Ladder Administrator/Supervisor Evaluation Instrument and Standards Review Panel
[December, 1984]

n. Counselor Competency Development Task Force
[Fall 1984]

o. Library Media Specialists Competency Development Task Force
[Fall 1984]

p. Counselor Standards Setting Panel
[March 19, 1986]

q. Library Media Specialists Standards Setting Panel
[March, 1986]

r. Elementary Counselor Standards Setting Panel
[January 30, 1987]

s. Speech/Language Standards Setting Panel
[January 21, 1988]

t. School Psychologists Standards Setting Panel
[January 21, 1988]

C. Identification of the domains, indicators, and measurement items

1. Review of research literature on effective teaching and effective schools

a. Interim Certification Commission Meeting, Proposed Criteria, April 4, 1983

b. References, effective teaching and effective schools

c. Reports prepared for Tennessee State Department of Education

Rosenholtz, Susan J. and Robinson, Sally. Effective Teachers. Nashville, TN: Vanderbilt University, January 15, 1983.

Rosenholtz, Susan J. and Robinson, Sally. Organizational Context of Effective Schools. Nashville, TN: Vanderbilt University, January 15, 1983.

Rosenholtz, Susan J., Smylie, Mark A., and French, Russell. Master Teacher Program Proposed Criteria for Teacher Selection and Evaluation. Preliminary draft. Nashville, TN, March 31, 1983.

Rosenholtz, Susan J. and Smylie, Mark A. Master Teacher Program, Suggested Criteria for Master Teacher Selection and Evaluation. Report to the Commissioner, Tennessee State Department of Education, Revised Draft, June 23, 1983.

Rosenholtz, Susan J. and Smylie, Mark A. Master Teacher Program, Effective Teaching Behaviors and Practices: A Synthesis of Classroom Research. Report to the Commissioner, Tennessee State Department of Education, Partial Draft, For Discussion Only, June 23, 1983.

Rosenholtz, Susan J. and Smylie, Mark A. Effective Teaching Strategies: A Synthesis of Classroom Research. Report to the Commissioner, Tennessee State Department of Education, Master Teacher Program, June, 1983, Revised December, 1983.

2. Review of additional research

   a. Evaluation

   b. Measurement

3. Review of evaluation plans in other states

   a. California Master Teacher Program (Sam Marino Unified School District)

   b. Charlotte-Mecklenburg Career Development Plan, North Carolina

   c. Georgia Performance Based Teacher Certification

   d. Houston Independent School District

   e. King William County Public Schools' Teacher Incentive Program, King William, Virginia

   f. School District of the City of Ladue Evaluation plan, St. Louis, Missouri

   g. New Albany Plan, New Albany, Mississippi

   h. Toledo Public Schools Evaluation Plan, Toledo, Ohio

   i. Virginia Beginning Teaching Assistance Program

4. Review of other evaluation information

   a. Alcoa Aluminum Evaluation Plan

   b. Evaluation Packet, Tennessee Education Association - Tennessee School Boards Association Joint Committee on Professional Personnel Evaluation

   c. Evaluation Workshop, Tennessee Education Association, October, 1983

   d. Xerox Evaluation Plan

5. Review of instate LEA evaluation plans

   a. Summary Analysis of LEA Performance Appraisal Plans, Summer 1983

   b. Meeting With School Systems With Laudatory Plans, September 1, 1983

   c. Key Elements Identified in the Evaluation Plans and Presentations, September 6, 1983

   d. Oak Ridge Evaluation Plan

6. Administrator Competencies Project

   a. Selected Bibliography and References

   b. Competency review

7. Feedback from practitioners

   a. Teachers Study Council
   August 1, 1983 and October 26, 1983

   b. Principals, Supervisors

   c. Superintendents

   d. Vocational Education Teachers

   e. Special Education Teachers

   f. Librarians/Media Specialists

   g. Special Groups - attendance supervisors, audiologists, school counselors, school psychologists, school social workers, speech/language pathologists

8. Criteria for inclusion

   a. Legal Requirements

      (1) Examination of in-service and professional development activities undertaken by the applicant [T.C.A. 49-5-5204(a)(4)]

      (2) Review of indicators of student progress, where applicable [T.C.A. 49-5-5204(b)(3)]

      (3) Proficiency tests of the teacher's knowledge [original draft of Master Teacher Act of 1983, but not in 1984 CERA]

      (4) Fair, unbiased and objective [T.C.A. 49-5-5103(4)]

      (5) Non-discriminatory on basis of race or sex [T.C.A. 49-5-5103(4)]

      (6) Other appropriate criteria [T.C.A. 49-5-5204(b)(5)]

   b. Based on research [promotes student achievement]

   c. Professional agreement [at least 70 percent of teachers agreed]

   d. Reasonable [easy to collect data]

   e. Measurable

   f. Do's and Don'ts of Master Teacher Evaluation System [Ad Hoc ICC, July 28, 1983]

9. Individuals making final determinations

   a. Staff

      (1) Teacher System

         Dr. Russell French - Vita
         1983-1987

         Dr. Carol Furtwengler - Vita
         1983-1986

         Dr. Joy McLarty - Vita
         1984-1987

         Dr. George Malo - Vita
         1983-present

      (2) Administrator/Supervisor System

         Dr. Fran Trusty - Vita
         1984-1987

         Dr. Mary Martin - Vita
         1985-1987

      (3) Consultant

         Dr. Ernie Rakow - Vita
         Memphis State University
         1986-present

   b. Ad Hoc Interim Certification Commission, 1983-84

   c. Summary of Interim Commission Activities, April 4, 1983 - October 31, 1983

   d. Official Actions Taken by the Ad Hoc Interim Commission, April, 1983 - March, 1984

   e. Interim Certification Commission, 1984-1986

   f. History of the Ad Hoc Interim Commission and the Interim Certification Commission, April 1983-January 1986

   g. Actions Taken by Interim Certification Commission 1984-85

   h. State Certification Commission, 1986-87, 1987-88

   i. State Board of Education, 1984, 1988

10. Final Report, Planning Grant to Develop Teacher Incentive Structures, September 1986

D. Qualification of evaluators

1. Evaluator selection process

a. Selection of field test evaluators, March 21 - May 1, 1984

(1) Field Test Evaluators for the Teacher Career Ladder Evaluation System, Memorandum to District Directors, February 20, 1984

(2) Field Testing for the Teacher Career Ladder Evaluation System, Memorandum to Superintendents/ Directors of Schools, February 21, 1984

(3) Application for Field Test Evaluators, Memorandum to Potential Field Test Evaluators, February 24, 1984

(4) Field Test Evaluators for the Teacher Career Ladder Evaluation System, Memorandum to Deans, Colleges of Education, February 24, 1984

(5) Evaluator Characteristics, February, 1984

(6) Field Test Evaluators: Selection Process, March, 1984

(7) Letters to Field Test Evaluator Applicants, March, 1984

(8) Field Testing, Memorandum, March 1, 1984.

b. Selection of teacher and administrator evaluators, 1984-85

(1) Selection of Teacher and Administrator Evaluators for 1984-85, Memorandum to ICC, March 14, 1984

(2) Evaluator Characteristics, March 1984

(3) Application for Evaluators, Memorandum to Potential Evaluators, April 2, 1984

(4) Applications Subcommittee - Agenda Item 1 [ICC Minutes, 5/1-2/84]

(5) Applications Subcommittee - Agenda Item 2 [ICC Minutes, 5/1-2/84]

(6) Application Procedures for Administrator/Supervisor Evaluators, Memorandum to Superintendents, Principals, etc., May 4, 1984

(7) Tennessee Career Ladder Application for Teacher Evaluator

(8) Teacher Evaluator Selection Committee [ICC Minutes, 7/20/84]

(9) Proposed Administrator/Supervisor Evaluator Selections [ICC Minutes, 7/20/84]

(10) Procedures for Identifying, Screening,and Selecting the Remaining Evaluators for the 1984-85 School Year [ICC Minutes, 8/19/84]

(11) Report of the Subcommittee of the Teacher Evaluator Selection Committee [ICC Minutes, 8/19/84]

(12) Percentage of Teachers and Percentage of Evaluators by Region [ICC Minutes, 8/19/84]

(13) Selection of Evaluators for Field Testing Special Education and Vocation Education Evaluation Systems [ICC Minutes, 1/11/85]

c. Selection of Evaluators, 1985-86

(1) Re-employment of Current Evaluators [ICC Minutes, 3/15/85]

(2) Procedures for Selecting 1985-86 Evaluators [ICC Minutes, 6/28/85]

d. Selection of Evaluators, 1986-87

(1) Evaluator Selection Procedures for 1986-87 [ICC Minutes, 4/8/86]

e. Selection of Evaluators, 1987-88

(1) Letter to Career Level II/Career Level III Teachers, March 16, 1987

(2) Plan for Use of Part-Time Evaluators in 1987-88 [SCC Minutes, 5/8/87]

(3) Selection of Teacher Evaluators [SCC Minutes, 6/19/87]

(4) Evaluator Job Description [SCC Minutes, 8/7/87]

f. Evaluator Application Form

2. Evaluator training process

a. Field test evaluator training, March 26 - April 4, 1984

(1) Training Program for Field Test Evaluators Memorandum to ICC, March 14, 1984

(2) Field Test Evaluator Training Program, Schedule of Events

(3) Evaluator Training Manual, Field Test, Spring 1984

(4) Update on Field Test Evaluator Training [ICC Minutes, 5/1-2/84]

b. Evaluator training, 1984-85

(1) Alternative Evaluator Training Model I
Alternative Evaluator Training Model II

(2) Career Ladder Evaluator Code of Ethics [ICC Minutes, 8/19/84]

(3) Evaluator Training Manual, 1984-85

c. Evaluator training, 1988-89

(1) Evaluator Training Manual, 1988-89

3. Evaluation of evaluators

a. Career Ladder Teacher Candidate Questionnaire, 1984-85

b. Career Ladder Administrator/Supervisor Evaluation System Candidate Questionnaire, 1985-86

c. Career Ladder Candidate Survey, 1985-86

d. Reliability checks during training and during evaluation process

(1) Career Ladder Teacher Evaluation Training, December 17-21, 1984, Agenda

(2) "Evaluator Training," Teacher Orientation Manual, 1986-87, p. 3

(3) "Evaluator Training and Reliability," Teacher Orientation Manual, 1987-88, p. 3

(4) "Evaluator Training and Reliability," Teacher Orientation Manual, 1988-89, p. 3

(5) Evaluator Training, Maxwell House, Fall 1988

e. 1987-88 Evaluator Training and Evaluation Procedures Memorandum to Dr. Brent Pulton, April 30, 1988

4. Qualifications and experience of evaluators

a. Field Test Evaluators, March 21 - May 1, 1984

(1) Persons Selected for Field Test Evaluator Training, March 19, 1984

b. Evaluators, 1984-85

(1) Tentative Teacher Evaluators, School Year 1984-85 (by Region) [ICC Minutes, 8/6/84, 8/19/84]

(2) Evaluators for Tennessee Administrator/Supervisor Evaluation System, 1984-85 [ICC Minutes, 7/20/84]

(3) Special Education, Vocation Teacher [Field Test] Evaluators, [ICC Minutes, 1/11/85, 1/26/85]

c. Evaluators, 1985-86

(1) Re-Employment of Special Education/Chapter I and Vocational Education Field Test Evaluators as Evaluators for 1985-86 [ICC Minutes, 6/28/85]

(2) Vocational Education, Special Populations Evaluators, 1985-86 [ICC Minutes, 7/24/85]

(3) Tennessee Career Ladder Administrator/Supervisor Evaluators for 1985-86 [ICC Minutes, 7/24/86]

(4) Report of the Evaluation Committee, Interim Certification Commission, August 2, 1985

(5) Report of the Evaluation Committee, Interim Certification Commission, August 7, 1985

d. Evaluators, 1986-87

(1) Recommendations for 1986-87 Evaluators: Teachers, Counselors, Librarians (for re-employment) [SCC Minutes, 7/2/86]

- (2) Teacher Evaluators, 1986-87 [SCC Minutes, 7/24/86, 8/19/86]
- (3) Additional Administrator/Supervisor Evaluators, 1986-87 [SCC Minutes, 8/19/86]
- (4) Teacher Evaluators, 1986-87 [SCC Minutes, 12/16/86]

e. Evaluators, 1987-88

- (1) Evaluators to be Employed in Administrator/Supervisor Evaluation System for 1987-88 [SCC Minutes, 6/19/87]
- (2) Teacher Evaluators, 1987-88 [SCC Minutes, 6/25/87, 8/7/87, 9/11/87]

f. Evaluators, 1988-89

- (1) Teacher Evaluators, 1988-89 [SCC Minutes, 7/7/88, 9/9/88]
- (2) Proposed 1988-89 Administrator/Supervisor Evaluators [SCC Minutes, 7/7/88]
- (3) Confirmation of Additional Part-time Administrator/ Supervisor Evaluators, 1988-89 [SCC Minutes, 7/7/88]

5. Assignment procedures for evaluators

a. Legal requirements

- (1) Team of properly trained Career Level III teachers or Career Level II (III) principals or supervisors from outside applicant's school system [T.C.A. 49-5-5205(b)(1)]
- (2) At least one evaluating teacher from same general grade area or subject area as applicant [T.C.A. 49-5-5205(g)]
- (3) Applicant may request that one member of team be removed and new member named by Commission [T.C.A. 49-5-5205(h)]

b. Scheduling Evaluations [SCC Minutes, 4/10/87]

c. Plan for Use of Part-time Evaluators in 1987-88 [SCC Minutes, 5/8/87]

d. Selection of Teacher Evaluators [SCC Minutes, 6/19/87]

e. Report on Evaluator Selection and Team Assignments

E. Feedback and improvement processes

1. Process for soliciting feedback and recommendations regarding desired improvements

   a. Revision process, 1985

      (1) Revisions Prompted by Dissemination of Rating Scales [ICC Minutes, 1/26/85, 2/20/85]

      (2) Evaluation of the 1984-85 Evaluation Process Review and Revision of Teacher Evaluation System: Description of Activities [ICC Minutes, 5/10/85]

      (3) Review and Revision of Career Ladder Teacher and Administrator Evaluation Systems [Career Ladder Retreat, April 22, 1985]

      (4) Letter and Questionnaire sent to Career Ladder candidates, May 21, 1985

      (5) Teacher Education Questionnaire [printed in Teacher Education, SDE Newsletter]

      (6) Teacher Evaluation System Review by Evaluators, Memorandum, July 5, 1985

      (7) Status Report, Review and Revision of the Teacher and Administrator Evaluation Systems [Career Ladder Retreat, July 8, 1985]

      (8) Recommendations made by Dr. Jason Millman Regarding Career Ladder Revisions, July, 1985

      (9) Revisions in Current Evaluation Systems Recommended for 1985-86 [ICC Minutes, 8/16/85]

   b. Revision process, 1986

      (1) State Certification Commission, Evaluation Committee Retreat, July 30-31, 1986

         - Teacher Candidate Surveys

         - Career Ladder Sounding Boards - General Education, Vocational Education, Special Education, Counselors, Librarians, Accelerated Career Development participants

         - Teacher Evaluator Feedback Dialogue

- Administrator/Supervisor Sounding Board
- Administrator/Supervisor Candidates
- Administrator/Supervisor Evaluators
- Statistical Analyses of Evaluation Scores
- Career Ladder Test Review Panel

(2) Recommended Improvements in Teacher Evaluation System for 1986-87 [SCC Minutes, 8/19/86]

(3) Recommended Improvements in Administrator/ Supervisor Evaluation System for 1986-87 [SCC Minutes, 8/19/86]

c. Revision process, 1987

(1) Tennessee Certification Commission, Evaluation Committee, March 27, 1987

(2) Review and Revision of Teacher Evaluation System [SCC Minutes, 4/10/87]

(3) Review and Revision of Teacher Evaluation System, Responses to Study Questions [SCC Minutes, 5/8/87]

(4) Update on Data Analysis of Career Ladder Evaluation System for Teachers Preliminary Report, 5/7/87

(5) Recommended Improvements in Administrator/ Supervisor Evaluation System [SCC Minutes, 5/8/87]

(6) Recommended Changes in Teacher Evaluation System for 1987-88 [SCC Minutes, 6/19/88]

(7) Recommended Improvements in Administrator/ Supervisor Evaluation System for 1987-88 [SCC Minutes, 6/19/88]

d. Revision process, 1988

(1) Review of Counselor Evaluation System Data [SCC Minutes, 1/21/88]

(2) Consideration of Addition of Student Progress as an Evaluating Factor [SCC Minutes, 1/21/88, 3/3/88, 5/5/88]

(3) Consideration of Inclusion of Field Tested Items in Teacher Evaluation System [SCC Minutes, 6/10/88]

(4) Consideration of Recommended Improvements in Administrator/Supervisor Evaluation System [SCC Minutes, 6/10/88]

(5) Approval of Weights Assigned to Field Tested Items in Teacher Evaluation System [SCC Minutes, 7/7/88]

(6) Approval of Weights for Principal/Assistant Principal Evaluation System, 1988-89 [SCC Minutes, 9/9/88]

(7) Discussion of Standards for Evaluation Systems [SCC Minutes, 9/9/88]

(8) Approval of Changes in Counselor Evaluation System [SCC Minutes, 11/4/88]

(9) Summary Report, Evaluator Debriefing, 3/22/88

(10) The Career Ladder Program, Summary

2. Feedback to persons providing input on desired improvements

    a. Communication Plan for Career Levels II and III, Memorandum, April 23, 1985

    b. Pacesetters, a professional update for educators moving up the Career Ladder

    c. Orientation Manuals

    d. State (Interim) Certification Commission Updates
    State Board of Education Updates

    e. Tennessee Education [State Department of Education newsletter]

    f. Memos to Superintendents, Principals, Teachers

3. Criteria used to select improvements to be made

    a. Are required by the law or any amendments to the law

    b. Are supported by data analysis

    c. Increase fairness and objectivity

    d. Delete non-differentiating items

    e. Eliminate repetitive items

    f. Simplify administration of the evaluation process

- g. Decrease paperwork for candidates
- h. Refine and clarify the evaluation process
- i. Can be accomplished without changing standards

4. Approval process for making changes
    - a. Staff recommendations
    - b. Consideration and approval by State (Interim) Certification Commission
    - c. Approval by Tennessee State Board of Education

II. Construct Validity

A. Multi-data source methods

1. Background on instrument development, April, 1983 - December 1983

   a. Evaluation Interviews, Teacher Portfolios [Ad Hoc Interim Certification Commission (ICC), 4/4/83]

   *b. Domains of Teacher Competence and Proposed Means of Assessment [Ad Hoc ICC, 6/23/88]

   *c. Flow Chart for Testing and Finalizing Evaluation Instruments and Developing Scales for Determination of Master Teacher Status

   d. Instruments for Review, Memorandum to Ad Hoc ICC, July 18, 1983

   *e. [Analysis of] Relationship Between Questionnaires and Master Teacher Competencies [Ad Hoc ICC, 7/19/83]

   *f. Tentative Components of Teacher Evaluation System, Master Teacher Program [Ad Hoc ICC, 9/15/83]

   *g. Relationships Between Teacher Competencies and Evaluation Instrument Items: Samples From Teaching Strategies Competencies, Rosenholtz and Smylie, Presented to Ad Hoc ICC, September 15, 1983

   *h. Relationship Between Teacher Competencies and Evaluation Instrument Items: Teaching Strategies Competencies, Rosenholtz and Smylie, Presented to Ad Hoc ICC, September 16, 1983

   *i. Analysis of Teacher Competencies, Sources of Data for Evaluation, and Weighting of Various Criteria, Memo to Ad Hoc ICC, October 13, 1983

   j. S.D.O.E. Staff Activities for Interim Commission, November-December, 1983 [Ad Hoc ICC, 12/15/83]

   *k. Competencies/Indicators and Instruments [Ad Hoc ICC, 12/15/83]

   l. Career Teacher Evaluation System, approved for field testing by Ad Hoc Interim Commission, December 14-15, 1983

m. Thirty Activities Required in the Development of Each Evaluation Component, December 1983

n. Meeting on Consultants and Contracts, February 1, 1984, Agenda and Memorandum, February 2, 1984

o. Sources of Data, Teacher Orientation Manual, 1984-85, pp. 19-38

2. Individual instrument development

a. Observation

(1) Observation Instruments for Teacher Evaluation, Memorandum to ICC, August 19, 1983

(2) Guidelines for Evaluating Teacher Observation Systems, Rosenholtz and Smylie, presented to ICC, September 15, 1983

(3) Approaches to Observation, Memorandum to ICC, October 31, 1983

(4) Sample Procedures for Observation, Memorandum to Teacher Study Council Representatives, November 1, 1983

*(5) The Stallings Observation System, November 8, 1983

*(6) Analysis of Classroom Management and Teaching Strategies, Competencies/Observation System, Memorandum, November 9, 1983

*(7) Observations Regarding SOS (Stallings Observation Model), Memorandum, November 9, 1983

(8) Meeting with Dr. Robert Soar, University of Florida, November 17 and 18, 1983, S.D.O.E. Activities for ICC [ICC Minutes, 12/15/83]

(9) Comments from Bob Soar, 11/17/83 - 11/18/83

(10) Classroom Observation Form, Draft Copy, Copyright 1984, State of Tennessee

b. Questionnaires: Student, peer, and principal (superordinate)

(1) Questionnaire for Elementary students
Questionnaire for Secondary Students, Rosenholtz and Smylie, presented to Ad Hoc ICC, September 15, 1983

*(2) Relationship Between Student Questionnaire Items and Master Teacher Competencies, Rosenholtz and Smylie, presented to Ad Hoc ICC, September 15, 1983

(3) Drafts of proposed items for student questionnaire with competencies and indicators prepared by staff, Fall 1983

(4) Student Questionnaires, Memorandum to ICC [office use only], October 26, 1983

(5) Legality of Using Anonymous Student Evaluations of Their Teachers, Memorandum, December 12, 1983

(6) Drafts of student questionnaires, elementary and secondary, prepared by staff, February 1984

(7) Drafts Student Questionnaires, approved for field testing, March, 1984

(8) Staff met with Anderson, Niebuhr & Associates, marketing research specialists, St. Paul, Minnesota, to review questionnaires, Spring 1984

(9) Drafts of proposed items for peer questionnaire with competencies and indicators prepared by staff, Fall 1983

(10) Draft of proposed peer questionnaire prepared by staff, February, 1984

(11) Drafts of proposed questions for structured interview, principal's evaluation of Master Teacher candidate, June 3, 1983

(12) Draft of Peer Questionnaire approved for field testing, March, 1984

(13) Draft of Superordinate Interview [Principal Questionnaire] approved for field testing, March, 1984

c. Candidate Interview

(1) Proposed Questions for Structured Interview, Master Teacher Review and Evaluation, June, 1983

(2) Structured Interviews for Teacher Candidates [Ad Hoc ICC, 10/20/83]

(3) Tennessee Career Ladder Evaluation System, Bibliography; Competencies/Indicators/Measures

(Interview Items) prepared by Appalachia Educational Laboratory, February, 1984

(4) Candidate Interview, Teacher Evaluation System, copyright 1984, approved for field testing, March, 1984

(5) Candidate Interview, Revised, April, 1984

d. Portfolio

(1) Proposed "Student Achievement" Questions, Memorandum to Members of the Application, Portfolio, Interview Subcommittees, Ad Hoc ICC, July 20, 1983

(2) Directions for Completion of Teacher Resume, approved by Ad Hoc ICC for field testing, December, 1983

*(3) Portfolio - Lesson Plans, Competencies/Indicators/ Evidence in the Portfolio

(4) Portfolio Materials, Memorandum to Field Test Evaluators, April 11, 1984

(5) Portfolio, Memorandum to Field Test Evaluators, April 19, 1984

(6) Draft of Portfolio Summary, approved for field testing, Spring 1984

(7) Evaluator Feedback on Portfolio, Spring 1984

(8) Teacher Portfolio, Rating and Summary, Revised, May, 1984

e. Career Ladder Test

(1) Request for Proposal, State of Tennessee, Department of Education, 1983-1984

- Vendors Sent Proposal
- IOX, Qualifications of Staff
- Measurement Incorporated, Qualifications of Staff
- Scholastic Testing Service, Qualifications of Staff
- Proposal Evaluation Sheet
- Evaluation Points

(2) Developing Four Assessment Instruments for the Tennessee Department of Education's Personnel Appraisal System for Tennessee Educators, Submitted by IOX Assessment Associates, 114111 West Jefferson Boulevard, Culver City, California, January 13, 1984

(3) Contract Between the State of Tennessee, Department of Education and IOX Assessment Associates, March 2, 1984

(4) Teachers to Participate in Field Testing the Written Test, Memorandum, March 23, 1984

3. Instrument pilot test

   a. Letter to Mr. Charles Frazier, Director, Metro/Davidson County Public Schools, February 23, 1984

4. Field test, April 2-27, 1984

   a. Sampling design, prepared by State Testing and Evaluation Center, University of Tennessee, Knoxville, February 7, 1984

   b. Field Testing, Proposed Schedule for Career Ladder For Teachers

   c. Field Testing for the Teacher Career Ladder Evaluation System, Memorandum to Superintendents/Directors of Schools, February 21, 1984

   d. Teacher Evaluation Program Consultant Services Needed, Training Program for Field Test Evaluators, Memorandum, February 23, 1984

   e. Categories for Grades 7 and 8, Categories for Grades 9-12

   f. Career Ladder Evaluation System Field Test, Memorandum to District Directors, March 1, 1984

   g. Participation in Field Testing of the Career Ladder Evaluation System, Memorandum to Superintendents/Directors of Schools, March 1, 1984

   h. Career Ladder "Candidate" Phone Script [Field Test]

   i. Schools and Grades/Subject Areas to be Sampled for Field Test [by Educational Developmental District]

j. Distribution of Field Test Participants by Subject Area (Grades 7-12)

k. Letter to Field Test Participant, March 26, 1984

l. Career Ladder Field Test Background Questionnaire, April, 1984

m. Letter to Field Test Candidates and Field Test Candidate Questionnaire, May, 1984

n. Follow-up letter to Field Test Candidates, July 5, 1984

o. Field Test Evaluator Feedback, June, 1984

p. Field Test Instruments

5. Instrument Approval

a. Teacher Evaluation System Standard Setting Panel [ICC Minutes, 8/6/84]

b. Summary of Recommendations from Panel on Career Ladder Instruments, August, 1984

c. Teacher Evaluation System Standard Setting Panel [ICC Minutes, 8/19/84]

d. Teacher System Instruments, 1984-85, 1985-86

6. Revisions in instruments

a. Revisions, 1985-86

(1) Proposals for Resolution of Problems Created by Release of Rating Scales [ICC Minutes, 2/8/85]

(2) Recommendations for Adjustment of Weights for the Teacher Evaluation System [ICC Minutes, 2/20/85]

(3) Revisions in Current Evaluation Systems Recommended for 1985-86 [ICC Minutes, 8/16/85]

b. Revisions, 1986-87

(1) Recommended Improvements in Teacher Evaluation System for 1986-87 [SCC Minutes, 8/19/86]

c. Revisions, 1987-88

(1) Tennessee Certification Commission, Evaluation Committee, March 27, 1987

   (2) Review and Revision of Teacher Evaluation System [SCC Minutes, 4/10/87]

   (3) Review and Revision of Teacher Evaluation System [SCC Minutes, 5/8/87]

   (4) Tennessee Certification Commission, Evaluation Committee, May 7, 1987 [SCC Minutes, 5/8/87]

   (5) Update on Data Analysis of Career Ladder Evaluation System for Teachers, May 7, 1987

   (6) Recommended Changes, Administrator Systems [SCC Minutes, 5/8/87]

   (7) Recommended Changes in Teacher Evaluation System for 1987-88 [SCC Minutes, 6/19/87]

   (8) Career Ladder Administrator/Supervisor Evaluation System, Recommended Improvements for 1987-88 [SCC Minutes, 6/19/87]

  d. Revisions, 1988-89

   (1) Consideration of Inclusion of Field Tested Item, Student Achievement, and Others in Teacher Evaluation System [SCC Minutes, 6/10/88]

7. Relationships among instruments and their implications regarding measurements of domains and indicators

  a. Items marked with an asterisk (*) under II-A.-1. and II-A.-2. above

  b. Update on Data Analysis of Career Ladder Evaluation System for Teachers, Preliminary Report, May 7, 1987

  c. Sources of Data, Teacher Orientation Manual, 1984-85

  d. Multiple Sources of Data, Teacher Orientation Manual, 1985-86

  e. Multiple Sources of Data, Teacher Orientation Manual, 1986-87

  f. Multiple Sources of Data, Teacher Orientation Manual, 1987-88

  g. Multiple Sources of Data, Teacher Orientation Manual, 1988-89

III. Criterion - Related Validity

A. Studies using existing data

1. Teacher of the Year District, State National Winners
2. TEA Distinguished Classroom Teachers
3. Tennessee National Candidates for Presidential Awards For Excellence in Science and Mathematics Teaching
4. Tennessee Foreign Language Teaching Association Outstanding Teachers of the Year
5. Awards Received by Members of Tennessee Association for Health, Physical Education, Recreation and Dance
6. Tennessee Teachers Selected as Semi-Finalists For NASA Teacher in Space Project

IV. Setting Standards, Reporting Results

A. Setting weights and standards

(1) Analysis of Teacher Competencies, Sources of Data for Evaluation, and Weighting of Various Criteria, Memorandum to Ad Hoc ICC, October 13, 1983

*(2) Recommendations and Alternatives for Scoring Procedures, presented to Ad Hoc ICC, December 15, 1983

(3) Weighting for Career Teacher Evaluation System, presented to Ad Hoc ICC, January 15, 1983

*(4) Scoring Procedures, Career Teacher Evaluation System, approved by Ad Hoc ICC Interim Commission, December 14-15, 1983, pp. 23-24

(5) Career Ladder Evaluation Instrument and Standards Review Panel, August 1-2, 1984

*(6) Career Ladder Teacher Evaluation Instrument and Standards Review Panel, Summary of Recommendations and Back-up Information, August, 1984

*(7) Career Ladder Teacher Evaluation Instrument and Standards Review Panel, Working Documentation, August, 1984

*(8) Teacher Evaluation System Standard Setting Panel [ICC Minutes, 8/19/84]

*(9) Criteria Analysis Committee Report (ICC Minutes, 9/7/84]

a. Use of Tennessee Professional Skills Test in Evaluation of Candidates for Career Levels II and III

b. Evaluation Procedures and Standards for Certification of Career Level II and Career Level III Teachers

(10) Proposal for Resolution of Problems Created by Release of Rating Scales [ICC Minutes, 2/8/85]

(11) Recommendations for Adjustment of Weights for the Teacher Evaluation System [ICC Minutes, 2/20/85]

*(12) Revisions in Current Evaluation Systems Recommended for 1985-86 [ICC Minutes, 8/16/85]

*(13) Deletion of Student Questionnaire Data for Selected Special Education Teachers [SCC Minutes, 7/2/86]

(14) Approval of Weights Assigned to Field Tested Items in Teacher Evaluation System [SCC Minutes, 7/7/88]

B. Scoring process

(1) Items marked with an asterisk (*) under IV.-A. above

(2) Score Documentation, 1984-85, 1985-86, 1986-87, 1987-88

(3) _Your Career Ladder Score Report_, A Guide to Interpreting Your Scores, Teacher Edition, 1986, 1987, 1988

(4) Clarification of Conversion Scores on Principal and Superordinate Questionnaires [ICC Minutes, 1/10/86]

(5) Standards for Certification, Career Levels II and III, _Teacher Orientation Manual_, 1984-85, pp. 39-42

(6) Scoring and Equating, _Teacher Orientation Manual_, 1985-86, pp. 48-56

(7) Scoring, _Teacher Orientation Manual_, 1986-87, pp. 49-58

(8) Scoring, _Teacher Orientation Manual_, 1987-88, pp. 53-60

(9) Scoring, _Teacher Orientation Manual_, 1988-89, pp. 67-76

C. Equating procedures

(1) Revisions in Current Evaluation Systems Recommended for 1985-86 [ICC Minutes, 8/16/85]

(2) Recommended Improvements in Teacher Evaluation System for 1986-87 [SCC Minutes, 8/19/86]

(3) Evaluation Committee Report, Elimination of K-2 Questionnaire [SCC Minutes, 4/10/87]

(4) Approval of Weights Assigned to Field Tested Items in Teacher Evaluation System [SCC Minutes, 7/7/88]

(5) Discussion of Standards for Evaluation Systems [SCC Minutes, 9/9/88]

D. Reporting procedures

1. Reporting results to candidates
    - a. Procedures for Handling Recommendations of Candidates to Career Ladder Status [ICC Minutes, 5/10/85]
    - b. Teacher System Notification Letters, June, 1985
    - c. Teacher System Notification Letters, June, 1986
    - d. Teacher System Notification Letters, 1987
2. Reporting results to the public and others
    - a. Press Release, June, 1985
    - b. Career Ladder Program Summary, June, 1985
    - c. Press Releases and Summaries by Educational Developmental Districts, June, 1985 [announced at 6 sites across the State. Legislators and Superintendents were invited to be present.]
    - d. Press Releases - 1986, 1987, 1988
    - e. Teacher Education, Tennessee Department of Education newsletter, July 1985 and July 1986
    - f. "Message from the Commissioner"
      State Certification Commission Updates
      State Board of Education Updates
      [Mailed to superintendents, principals, supervisors, deans of colleges of education, presidents of Teachers Study Councils, evaluators, and statewide educational organizations (TEA, TOSS, TSBA)]
3. Interpreting results for candidates
    - a. Summary Conferences
    - b. Individual File Review
    - c. Career Ladder Technical Manual, A Guide to Interpreting Your Scores, Teacher Edition, 1986, 1987, 1988

V. Implementation of Fair Procedures

A. Definition of administrative procedures

1. Procedures for administering the evaluation
    - a. Evaluator Training Manual (See I-D-2)
    - b. Organization and Procedures of the State Certification Commission, Revised 5/5/88
    - c. Career Ladder Rules, Adopted by the State Board of Education
    - d. Legal Opinions
2. Guidelines provided evaluators
    - a. Evaluator's Handbook, 1986-87
    - b. Evaluator's Handbook, 1987-88
    - c. Evaluator's Handbook, 1988-89
    - d. "Messenger," evaluator newsletter
    - e. Memoranda to Evaluators
3. Information provided candidates
    - a. The Career Ladder Plan, letter sent to candidates, May 29, 1984
    - b. "Career Ladder Packet," sent to all teachers, June 4, 1984
    - c. "Portfolio In Brief," pamphlet provided candidates, Fall 1984
    - d. Application Form and sample receipt of application letter, 1988-89
    - e. Sample change or correction in application letter, 1988-89
    - f. Letter to Career Ladder Candidate regarding orientation sessions and Career Ladder Orientation Sites and Schedule, August 25, 1988

g. *Teacher Orientation Manual* (must have manual at least two weeks before Evaluation Visit "A")

h. Career Ladder State Test Study Guide

i. Memoranda to Candidates

j. Career Ladder Communication Plans

k. Hot Line Questions, sample

4. Information provided others

    a. Superintendents' Regional Drive-In Conference, June 12-13, 1984

    b. Career Ladder Program, Policies and Guidelines, Superintendents' Study Council, September 16-18, 1984

    c. Career Ladder Principal's Handbook, 1985-86, 1986-87, 1987-88, 1988-89

    d. State Certification Commission Updates

    e. *Tennessee Education*, newsletter from Tennessee Department of Education

    f. Memoranda to Superintendents, Directors, Principals

    g. Memoranda to Legislators, Higher Education, etc.

    h. Press releases

B. Implementation of administrative procedures

1. Degree to which procedures were carried out

    a. Career Ladder Test Administration

    b. Release of rating scales

2. Handling non-compliance with procedures

    a. Career Ladder Test Re-administered

3. Qualifications of individuals handling procedural problems

C. Changes in administrative procedures

1. Modifications in testing procedures
    - a. August 11, 1984, Administration of Tennessee Career Ladder Test: Problems and Recommended Solutions [ICC Minutes, 8/19/84]
    - b. Guidelines for Career Ladder Test Administration [ICC Minutes, 9/7/84]
    - c. Releasing Results of the Tennessee Career Ladder Test [ICC Minutes, 9/7/84]
    - d. Substitution of NTE Sub-Tests for Career Ladder Sub-Tests [ICC Minutes, 10/10/84]
    - e. Retaking the Professional Skills Portion of the Career Ladder Test [ICC Minutes, 10/10/84]
    - f. Upper Career Ladder Candidates Who Have Not Taken the Career Ladder Test During 1984-85 [SCC Minutes 6/28/85]
    - g. Upper Career Ladder Candidates Who Have Not Taken the Career Ladder Test During 1985-86 [SCC Minutes 8/19/86]
    - h. Approval of "Screening Procedures" for 1987-88 Career Ladder Evaluations [SCC Minutes, 10/10/86]
    - i. Candidates Who Have Not Taken the Career Ladder Test, 1986-87 [SCC Minutes 4/10/87]
    - j. Clarification of Completion of Test Requirements Prior to Evaluation [SCC Minutes, 4/10/87]
    - k. Special Problem: 1985-86 Non-Test Takers [SCC Minutes, 6/19/87]
2. Modifications in evaluation procedures
    - a. Review of Distribution of Rating Scales and Proposal for Resolution of Problems Created by Release of Rating Scales [ICC Minutes, 2/8/85]
    - b. Accelerated Career Development Program, 1985-86
        - (1) Career Development Opportunity for Career Ladder Candidates [ICC Minutes, 6/28/85]
        - (2) The Accelerated Career Development Program [ICC Minutes, 7/31/85]

(3) Update on Accelerated Career Development Program [ICC Minutes, 8/16/865]

(4) Update on Accelerated Career Development Program [ICC Minutes, 10/11/85]

c. Candidates Missing Evaluation Data [ICC Minutes 6/28/85]

d. Clarification of Status of Educators Evaluated with Insufficient Experience [ICC Minutes, 6/28/85]

e. Accelerated Career Development Programs for 1986-87 [SCC Minutes, 5/9/86]

f. Career Ladder Eligibility for Educators Evaluated in 1986-87 with Insufficient Experience [ICC Minutes, 2/18/87]

g. Accelerated Career Development Manual, 1985-86

3. Modifications in application procedures

a. Deadlines for Changes in Application Level in 1984-85 [ICC Minutes, 10/10/84]

b. Determination of Candidate's Eligibility [SCC Minutes, 1/30/87]

c. Consideration of Application Policy [SCC Minutes, 9/11/87]

D. Appeal procedures

1. Mandated appeal procedures

a. T.C.A. 49-5-5009 [1984 Supplement]

b. T.C.A. 49-5-5009 [1988 Supplement]

2. State Board of Education rules

a. 0520-2-2-.29 State Board of Education Career Ladder Certification Appeal Process

b. Memo from Executive Director of State Board, 3/27/89

3. State Certification Commission policies

a. Notification of Non-Recommendation [ICC Minutes, 3/15/85]

b. Commission Reactions to Proposed Appeals Decision [ICC Minutes, 1/10/86]

c. State Certification Commission's Response to Appeals Recommendations [SCC Minutes, 4/8/86]

d. Recommended Changes in Appeals Procedures [SCC Minutes, 5/9/86]

e. Discussion of Appeals Related Issues [SCC Minutes, 1/30/87]

f. Policy Regarding Appeals After Completion of ACD Program [SCC Minutes, 5/8/87]

4. Number of appeals

a. Update on Appeals Process [ICC Minutes, 8/16/85]

b. Report on Career Ladder Appeal Cases [SCC Minutes, 11/9/87]

c. Three-year Summary of Appeal Recommendations, 1984-85, 1985-86, 1986-87 [Fall 1988]

5. Individuals responsible for appeals

a. Paper review

(1) David Alexander, Appeals Administrator, State Board of Education

(2) Charles Ray, Appeals Administrator, State Board of Education

b. Career Ladder hearings

(1) The Honorable Jack Derryberry, Administrative Judge, State Board of Education, [1985-87]

(2) The Honorable Margaret Robertson - Vita Administrative Judge, State Board of Education, [1987-present]

Appendix D: Proposal - Tennessee Career Ladder System Validation Study

WORKING DRAFT
Proposal: Tennesse CLS Validation Study
(Revised August, 1989)

Purpose

The purpose of this document is to propose a plan for studying the validity of Tennessee's Career Ladder System (CLS). The proposed plan provides a description of the purpose and tasks of the study, a workplan, and a comprehensive list of anticipated data needs.

As discussed at the planning meeting held January 24-25, 1989 in Nashville, the general purpose of the validation study is to determine the extent to which the TN Career Ladder System (CLS) accurately identifies and differentiates among "good," "better," and "best" teachers (i.e., teachers who are assigned CLS attainment levels of 1 vs. 2 vs. 3). The January planning meeting also resulted in the conclusion that Year 1 of the study should depend on existing or readily available data and focus primarily on content- and criterion-related validity issues. Additional data collection needs and construct validity issues may be undertaken during a second year of study.

Tasks

The tasks listed below are proposed for Year 1 of the the CLS validation study:

1. Examine the relationship between the CLS (which consists of instruments, training procedures, implementation procedures, scoring, and the assignment of standards-based attainment levels) and the theoretical, empirical, and practical bases that support the CLS.

2. Examine the relationship between teachers' participation in the CLS and certain teacher charactistics (e.g., race, sex, educational background). The relationship between teachers' participation in the CLS and various teaching contexts (e.g., subject matter area, teaching assignment, class/school size, school location, students' test scores) will also be examined. Likewise, the relationships among teacher characteristics, teaching contexts, and levels of CLS attainment will be determined. These relationships will be studied over time in order to identify shifts in participation and attainment patterns based on teacher characteristics and teaching contexts.

3. Determine the extent to which scoring and weighting procedures accurately reflect levels of teaching excellence. This will be accomplished by considering various sources of content validity evidence and by examining the ways in which scores and attainment levels are influenced by (a) weighting within instruments and procedures and (b) relative weighting among and across the various instruments and procedures.

4. Compare CLS data with other available indicators of teaching excellence. The purpose of this comparison is to provide evidence of criterion validity.

5. Make recommendations concerning necessary and desirable areas for CLS change, continuing study, and/or additional study.

TASK 1 is concerned with *content validity*. TASK 2 can address issues of *equity/equal access and bias* (as they contribute to *construct validity*) by showing whether the system inappropriately differentiates among teachers on the basis of variables or factors other than teaching excellence. Analyses based on different points in time should help identify system modifications or other characteristics that have influenced changes in participation and attainment throughout the life of the system. TASK 3 can provide evidence for both *content and criterion-related validity*; TASK 4 is designed to provide evidence of *criterion-related validity*; and TASK 5 is expected to result in a plan for additional *construct validation* procedures.

Workplan

While the workplans for the tasks associated with criterion-related validity are implicit in the task descriptions above, the content validity tasks warrant more explicit attention, here. Hence, the following workplan particularly focuses on the content validity tasks. Throughout the Workplan description, "training" applies to both evaluators and evaluatees. "Instruments" include the following:

- Classroom Observation
- Teacher/Evaluator Dialogues (Planning Focus, Teaching Strategies Focus, and Evaluation Focus)
- Elementary Student, Secondary Student, and Principal Questionnaires
- Professional Development and Leadership Summary

The Workplan consists of addressing the four questions listed below with regard the instruments and related procedures. The questions have been designed to consider the "inputs," "processes," and "outcomes" of the CLS and its components (i.e., instruments, procedures). Input refers to the bases for the CLS and its components (e.g., identified needs, empirical evidence); process refers to developmental procedures; outcome refers to products and results (e.g., instruments, scores).

Question #1

What evidence exists that, as a whole, instrument development procedures reflected (a) the stated purposes of the CLS and (b) reasonable practice? Relevant data may include the following:

- Reports and documents of similar activities (e.g., teacher incentive programs in other states) that may be used for comparison purposes
- CLS legislation
- Evidence of consultation with experts
- Relevant literature reviews
- Evidence that equity/bias issues were considered throughout the process
- Needs assessment data
- Methodology descriptions

Question #2

What evidence exists that the content of each of the instruments reflects specified needs, reasonable practice, and empirical findings? Relevant data may include the following:

- Needs assessment data
- Evidence of consultation with experts
- Professional consensus building data
- Relevant literature reviews

Question #3

Overall, to what extent do training procedures ensure that the instruments are reliably used as intended? Relevant data may include the following:

- Information on the selection and training of trainers
- Information on the selection of trainees
- Information/observations regarding the depth, breadth, and duration of training
- Rater reliability data
- Information on retraining criteria and procedures
- Observations of training/retraining procedures
- Information on development/application of the certification standards
- Evidence of the extent to which training activities match actual practice

Question #4

Overall, to what extent are the instruments and related procedures (e.g., scoring) implemented as intended? Relevant data may include the following:

- Evidence of drift checks for rater reliability
- Evidence of objective scoring
- Anecdotal records
- Implementation evaluation data

Task Assignments

Members of the content validation team will address the four questions listed above as they apply to the Observations, Dialogues, Questionnaires, and the Professional Development and Leadership Summary. Team members will share primary responsibility for the instruments as follows:

- Observations--P. Sandifer
- Dialogues--T. Kerins
- Questionnaires & Professional Development/Leadership Summary--D. Redfield

Timeline

1/24 & 1/25	Team meeting with TN State Department Staff in Nashville
3/89	Draft proposal circulated among Team members for reaction
3/27/89	Team meeting with TN State Department Staff at AERA
3/29/89	Content validation team follow-up meeting
5/89	Develop and circulate lists of anticipated data needs among team members in preparation for June ECS meeting in Boulder
6/89	Meeting of available Team members and TN State Department Staff at ECS
8/89 - 10/89	Team members attend training sessions and work with files in Nashville
1/90	Draft report

Consolidated List of Anticipated Data Needs

The following data needs are listed by categories (e.g., those needs that apply to the "Overall" study, those that concern instrument development processes, . . . , numerical data required for various statistical analyses)

Overall

A statement of the purpose of the validation study.

Any documents that can provide an historical overview of the TN CLS from conceptualization to date.

Instrument Development Processes/Procedures

A copy of any TN CLS legislation (e.g., authorization, statement of purpose) and any related information that further defines or illustrates the intent of the legislation.

Descriptions of instrument development and modification processes, including descriptions/credentials of persons involved in various aspects of the processes.

Literature that informed or resulted from any of the instrument development-related processes (e.g., research reports, literature reviews). These could be in the areas of existing/emerging teacher evaluation systems, measurement theory, indicators of teacher effectiveness, court case outcomes, etc.).

Descriptions of any concensus gathering procedures (e.g., opinion leader surveys, focus interviews) and the persons involved (e.g., survey respondents, interviewers, survey developers).

Descriptions/credentials of consultants used throughout the instrument development processes.

Descriptions of any other information gathering and decision-making processes that may have influenced instrument development.

Instrument Content

Copies of the instruments currently in use.

Data bases that served as sources of information for the content pool (e.g., literature reviews, needs assessments, expert opinions, surveys and survey results, compendiums of instruments used in other teacher evaluation programs/states/districts).

Training

Training tapes

Observations of training sessions

Copies of training materials or related instructions.

Descriptions/documentation of training procedures, including:

- selection and training of trainers
- selection of evaluators (i.e., trainees)
- procedures for certifying trained evaluators
- establishment of instrument, interrater and/or intrarater reliability
- criteria for re-training and re-training provisions
- training conditions/logistics (e.g., duration, scheduling, location)
- training procedures
- training content

Implementation

Description/documentation of any follow-ups to training (e.g., observations of evaluators, drift checks).

Any other data that may be used to document instrument implementation (e.g., instructions for administration, self-reports such as interviews or surveys of users or recipients, annecdotal records).

Statistical Data Needs

a. Characteristics of participants:

- Gender
- Degree status
- School district
- Years of teaching experience
- Elementary vs. secondary
- Region of state
- Race/ethnicity
- Urban-rural classification of schools (cannot be provided because of the variety of school systems in TN, i.e., county, city, special school districts, and consolidated)

b. Internal analyses

(E. Rakow will redo the analyses in the May, 1987 report by French, Malo, and Rakow for the 1987-88 participants. Other specific analyses will be provided upon request)

c. Relationships to external factors

(TN State Department Staff are determining CLS participation and level attained by several groups of teachers, e.g., Teacher-of-the-Year finalists, winners of TEA's Distinguished Classroom Teacher Award, recipients of the Presidential Awards for Excellence in Science and Mathematics Teaching, winners of the Foreign Language Teacher-of-the-Year Award, and recipients of awards in Health and P.E.)

Appendix E: Validation Team Organizational Meeting Agenda

Agenda
Tennessee Career Ladder Validation Study
Planning Meeting
Nashville, Tennessee
January 24-25, 1989

Location:* Room 142, Cordell Hull Building

Time: 9:00 a.m., CST

I.	Purpose of Validation Study	George Malo
II.	Overview of Career Ladder Program	George Malo Judy Bassham
III.	Proposed Activities to Validate .... the Career Ladder Evaluation System	George Malo
IV.	Outline of Supporting Material	Betty Long
V.	Design of Validation Study	Validation Team

A. Role of Members

B. Timelines

C. Elements of Study

D. Procedures

* On January 25, the meeting will be held in Room 200, Cordell Hull Building.

Appendix F: Career Ladder System Instruments

PRE/POST-OBSERVATION CONFERENCE RECORD

Candidate: ____________________ S.S.#: ____-___-_____ Date: _______

Grade Level(s): ____/_____ Subject Area(s): ____________/__________

Evaluator/#: _______________/_______ Times:_____/_____ Visit: 1 2 3

PRE-CONFERENCE RECORD

POST-CONFERENCE RECORD

I. GENERAL NATURE/STUDENT OBJECTIVES

Obs. 1 ______________________________

/______________________________________

____

_______________ New ____ Review ____

Obs. 2 ______________________________

/______________________________________

____

_______________ New ____ Review ____

Was this congruent?

Yes No

Comments: _________________

____

____

Yes No

Comments: _________________

____

____

II. METHOD OF PRESENTATION

Obs. 1	Obs. 2		Obs. 1	Obs. 2
_____	_____	Lecture	_____	_____
_____	_____	Discussion/Interaction	_____	_____
_____	_____	Independent/Laboratory work	_____	_____

III. TEACHER ACTIVITIES

Obs. 1 ______________________________

/______________________________________

____

_______________ New ____ Review ____

Was this congruent?

Yes No

Comments: _________________

____

____

Obs. 2 ________________________ Yes No

/__________________________ Comments: ______________

_____________________________ ______________________

____________ New ___ Review ___ ______________________

IV. STUDENT ACTIVITIES

Obs. 1 ________________________ Was this congruent?
Yes No

/__________________________ Comments: ______________

_____________________________ ______________________

____________ New ___ Review ___ ______________________

Obs. 2 ________________________ Yes No

/__________________________ Comments: ______________

_____________________________ ______________________

____________ New ___ Review ___ ______________________

V. MEASUREMENT OF ACHIEVEMENT

Obs. 1 ________________________ Was this congruent?

_____________________________ Yes No

Obs. 2 ________________________ Comments: ______________

____

VI. RELATIONSHIP OF LESSONS

Yesterday's lesson: Tomorrow's lesson:

Obs. 1 ________________________ ______________________

_____________________________ ______________________

Obs. 2 ________________________ ______________________

_____________________________ ______________________

VII. SPECIAL CODES

___ unannounced observation
___ mainstreamed learners
___ special group: gifted
___ inappropriate facilities
___ organizational disfunction
___ special education: code # ______________________________

___ low socio-economic status
___ special group: slow learners
___ lack of materials
___ inappropriate grade
___ learners with special problems

General Comments:

Additional Evaluator Comments:

Teacher Comments:

Printed in U.S.A. NCS Trans-Optic® MP18-75035-32 A1407

TENNESSEE CAREER LADDER
CLASSROOM OBSERVATION

ED-2510
REV. July 1987

(PLEASE PRINT)
CANDIDATE
NAME: ______________ LAST ______________ FIRST ____ MI

SYSTEM: ______________

OBSERVATION DATE: ___/___/___ M D Y

SCHOOL: ______________

EVALUATOR
NAME ______________ LAST ______________ FIRST ____ MI

VISIT:	A1 ○	B1 ○	C1 ○	X1 ○
	A2 ○	B2 ○	C2 ○	X2 ○

GROUP:	GENERAL ED ○ SPECIAL ED (GROUP 1) ○ VOCATIONAL ED ○ LIBRARY MEDIA SPECIALIST ○	CHAPTER 1 ○ OTHER ○ SPECIFY ______________

MARK IF ANY OF THE FOLLOWING CONDITIONS EXIST.

○ UNANNOUNCED OBSERVATION

○ LOW SOCIO-ECONOMIC STATUS

○ ONE OR MORE MAINSTREAMED LEARNERS

○ SPECIAL GROUP, SLOW LEARNER

○ SPECIAL GROUP, GIFTED

○ LACK OF INSTRUCTIONAL MATERIALS

○ ASSIGNED TO INAPPROPRIATE FACILITIES

○ ASSIGNED INAPPROPRIATE COURSE/GRADE

○ ORGANIZATIONAL DYSFUNCTION

○ ONE OR MORE LEARNERS WITH SPECIAL PROBLEMS; e.g. PHYSICAL DISORDER, EMOTIONAL PROBLEM

EVALUATOR NUMBER		
⓪	⓪	⓪
①	①	①
②	②	②
③	③	③
④	④	④
⑤	⑤	⑤
⑥	⑥	⑥
⑦	⑦	⑦
⑧	⑧	⑧
⑨	⑨	⑨

CANDIDATE NUMBER								
⓪	⓪	⓪	⓪	⓪	⓪	⓪	⓪	⓪
①	①	①	①	①	①	①	①	①
②	②	②	②	②	②	②	②	②
③	③	③	③	③	③	③	③	③
④	④	④	④	④	④	④	④	④
⑤	⑤	⑤	⑤	⑤	⑤	⑤	⑤	⑤
⑥	⑥	⑥	⑥	⑥	⑥	⑥	⑥	⑥
⑦	⑦	⑦	⑦	⑦	⑦	⑦	⑦	⑦
⑧	⑧	⑧	⑧	⑧	⑧	⑧	⑧	⑧
⑨	⑨	⑨	⑨	⑨	⑨	⑨	⑨	⑨

GRADE LEVEL OF STUDENTS
Ⓚ
①
②
③
④
⑤
⑥
⑦
⑧
⑨
⑩
⑪
⑫

CLASS START TIME			
⓪	⓪	⓪	⓪
①	①	①	①
②	②	②	②
③	③	③	③
④	④	④	④
⑤	⑤	⑤	⑤
⑥	⑥	⑥	⑥
⑦	⑦	⑦	⑦
⑧	⑧	⑧	⑧
⑨	⑨	⑨	⑨

SUBJECT CODE	
⓪	⓪
①	①
②	②
③	③
④	④
⑤	⑤
⑥	⑥
⑦	⑦
⑧	⑧
⑨	⑨

NUMBER OF STUDENTS	
⓪	⓪
①	①
②	②
③	③
④	④
⑤	⑤
⑥	⑥
⑦	⑦
⑧	⑧
⑨	⑨

NUMBER OF ADULTS	
⓪	⓪
①	①
②	②
③	③
④	④
⑤	⑤
⑥	⑥
⑦	⑦
⑧	⑧
⑨	⑨

SPECIAL EDUCATION

CONDITIONS		OPTIONS
⓪	⓪	①
①	①	②
②	②	③
③	③	④
④	④	⑤
⑤	⑤	⑥
⑥	⑥	⑦
⑦	⑦	⑧
⑧	⑧	⑨
⑨	⑨	⑩

○ RELIABILITY CHECK FOR

EVALUATOR NUMBER		
⓪	⓪	⓪
①	①	①
②	②	②
③	③	③
④	④	④
⑤	⑤	⑤
⑥	⑥	⑥
⑦	⑦	⑦
⑧	⑧	⑧
⑨	⑨	⑨

© 1987, Career Ladder Program, Tennessee State Department of Education

SECTION 1

OBSERVATION NO. 1

		Teacher Behavior	Positive Affect	Student Initiates/Responds
TASK	Overview	○	○	○
	Directions	○	○	○
	Examples	○	○	○
	Distinguishes Steps	○	○	○
CONTENT	Definition/Examples	○	○	○
	Explanations/Rules/Attributes	○	○	○
	Provides Context	○	○	○
	Uses Learner Remarks	○	○	○
REVIEW	Definition/Examples	○	○	○
	Explanations/Rules/Attributes	○	○	○
	Provides Context	○	○	○
	Uses Learner Remarks	○	○	○
PRACTICE	Initial Practice	○	○	○
	Independent Practice	○	○	○
	Varied Activities/Materials	○	○	○
	Group Practice	○	○	○
MONITORS	Requests Affirmation	○	○	○
	Requests Infor./Explanation	○	○	○
	Requests Comparison/Contrast	○	○	○
	Requests Evaluation	○	○	○
	Requests Questions	○	○	○
	Monitors Individual/Group	○	○	○
FEEDBACK	Recognizes Response	○	○	○
	Encourages Response	○	○	○
	Ignores Response	○	○	○
	Recognizes Correct Response	○	○	○
	Corrective Statement	○	○	○
	Corrected Assignments/Tests	○	○	○
	Reports Student Status	○	○	○
	Provides Recommendation	○	○	○
OTHER	Classroom Procedure	○	○	○
	Presentation/Other	○	○	○
	Teacher Off Task	○	○	○
	Greeting/Socialization	○	○	○

OBSERVATION BEGAN AT ____________

OBSERVATION SCHEDULED TO END AT ____________

OBSERVATION ENDED AT ____________

TIME ELAPSED BETWEEN THE TIME THE INTERRUPTION ENDED AND THE CLASS WAS BACK ON TASK IN MINUTES (M) AND SECONDS (S).

M	S	S	M	S	S	M	S	S
⓪	⓪	⓪	⓪	⓪	⓪	⓪	⓪	⓪
①	①	①	①	①	①	①	①	①
②	②	②	②	②	②	②	②	②
③	③	③	③	③	③	③	③	③
④	④	④	④	④	④	④	④	④
	⑤	⑤		⑤	⑤		⑤	⑤
	⑥	⑥		⑥	⑥		⑥	⑥
	⑦	⑦		⑦	⑦		⑦	⑦
	⑧	⑧		⑧	⑧		⑧	⑧
	⑨	⑨		⑨	⑨		⑨	⑨

TEACHER INVOLVES STUDENTS WHO DO NOT VOLUNTEER

YES ○ NO ○

TEACHER CONTROL

CAUSE						TEACHER ACTION	RESULT						
STUDENT BEHAVIOR							STUDENT BEHAVIOR						
OFF TASK			DISRUPTIVE				COMPLIANCE	OFF TASK			DISRUPTIVE		
INDIVIDUAL	SMALL GROUP	LARGE GROUP	INDIVIDUAL	SMALL GROUP	LARGE GROUP			INDIVIDUAL	SMALL GROUP	LARGE GROUP	INDIVIDUAL	SMALL GROUP	LARGE GROUP
○	○	○	○	○	○	○	○	○	○	○	○	○	○
○	○	○	○	○	○	○	○	○	○	○	○	○	○
○	○	○	○	○	○	○	○	○	○	○	○	○	○
○	○	○	○	○	○	○	○	○	○	○	○	○	○
○	○	○	○	○	○	○	○	○	○	○	○	○	○
○	○	○	○	○	○	○	○	○	○	○	○	○	○
○	○	○	○	○	○	○	○	○	○	○	○	○	○
○	○	○	○	○	○	○	○	○	○	○	○	○	○
○	○	○	○	○	○	○	○	○	○	○	○	○	○
○	○	○	○	○	○	○	○	○	○	○	○	○	○

ACADEMIC		PROCEDURAL	
⓪	⓪	⓪	⓪
①	①	①	①
②	②	②	②
③	③	③	③
④	④	④	④
⑤	⑤	⑤	⑤
⑥	⑥	⑥	⑥
⑦	⑦	⑦	⑦
⑧	⑧	⑧	⑧
⑨	⑨	⑨	⑨

ACADEMIC ____________

PROCEDURAL ____________

SECTION 2

OBSERVATION NO. 9

1. How many students are:

IN CLASS		OFF TASK	
⓪	⓪	⓪	⓪
①	①	①	①
②	②	②	②
③	③	③	③
④	④	④	④
⑤	⑤	⑤	⑤
⑥	⑥	⑥	⑥
⑦	⑦	⑦	⑦
⑧	⑧	⑧	⑧
⑨	⑨	⑨	⑨

2. During most of the observation period, the teaching method was
   - Lecture ○
   - Discussion/Interaction ○
   - Independent Work ○

3. During most of the observation period, the class was working:
   - As a total group ○
   - As subgroups ○
   - As individuals ○

4. During the observation period, did students express/exhibit lack of understanding?
   - None ○
   - One ○
   - Few (2-5) ○
   - Many (more than 5) ○

5. Did the teacher respond to this by reteaching?
   - Yes ○
   - No ○
   - N/A ○

6. Did the teacher present correct information?
   - Yes ○
   - No ○
   - N/A ○

7. Were there any safety/sanitation violations?
   - Yes ○
   - No ○
   - N/A ○

8. Did the teacher provide reasonable sanctions for inappropriate behavior?
   - Yes ○
   - No ○
   - N/A ○

DO NOT WRITE OUTSIDE BOXED IN AREA

COMMENTS

9. MATERIALS/ EQUIPMENT	USE	UNRELATED
VISUAL	○	○
KINESTHETIC/TACTILE	○	○
AUDITORY	○	○
REALIA/MODELS	○	○
PRINT	○	○
OLFACTORY	○	○
INTERACTIVE	○	○

SECTION 3

PLANNING/STRATEGIES/ MANAGEMENT	TEACHER PREPARATION	①	②	③	④	⑤	ⓧ
	STUDENT PREPARATION	①	②	③	④	⑤	ⓧ
	TEACHER ATTENTION TO NEEDS IN ORDER OF IMPORTANCE	①	②	③	④	⑤	ⓧ
	TEACHER PACING OF ACTIVITIES TO ACCOMMODATE LEARNER UNDERSTANDING	①	②	③	④	⑤	ⓧ
	TEACHER DEVELOPS HIGHER LEVEL THINKING SKILLS	①	②	③	④	⑤	ⓧ
	TEACHER ENCOURAGES STUDENT PARTICIPATION	①	②	③	④	⑤	ⓧ
USE OF RESOURCES	ACCESSIBILITY OF MATERIALS/MEDIA TO STUDENTS	①	②	③	④	⑤	ⓧ
	EFFECTIVE USE OF MATERIALS/MEDIA/RESOURCES BY TEACHER	①	②	③	④	⑤	ⓧ
	MATERIALS/RESOURCES APPROPRIATE TO OBJECTIVES	①	②	③	④	⑤	ⓧ
	MATERIALS/RESOURCES APPROPRIATE TO LEARNERS	①	②	③	④	⑤	ⓧ
	MAKES EFFECTIVE USE OF TIME	①	②	③	④	⑤	ⓧ
FACILITIES	MAKES EFFECTIVE USE OF AVAILABLE FURNITURE/EQUIPMENT	①	②	③	④	⑤	ⓧ
	MAKES EFFECTIVE USE OF AVAILABLE DISPLAY SPACE	①	②	③	④	⑤	ⓧ
TEACHER FEEDBACK	USE OF FEEDBACK	①	②	③	④	⑤	ⓧ
	USE OF EVALUATION	①	②	③	④	⑤	ⓧ
COMM	PROVIDES AN APPROPRIATE LANGUAGE MODEL FOR LEARNER	①	②	③	④	⑤	ⓧ
VOC EDUC	RELATES LESSON TO JOB SITUATIONS	①	②	③	④	⑤	ⓧ
CLIMATE	ESTABLISHES ACADEMIC ENVIRONMENT CONDUCIVE TO LEARNING	①	②	③	④	⑤	ⓧ
	DEMONSTRATES APPRECIATION FOR LEARNER EFFORTS/ACCOMPLISHMENTS	①	②	③	④	⑤	ⓧ
		①	②	③	④	⑤	ⓧ
		①	②	③	④	⑤	ⓧ
		①	②	③	④	⑤	ⓧ
		①	②	③	④	⑤	ⓧ
		①	②	③	④	⑤	ⓧ
		①	②	③	④	⑤	ⓧ
		①	②	③	④	⑤	ⓧ
		①	②	③	④	⑤	ⓧ

FAVORITISM

DID THE TEACHER SHOW FAVORITISM TO STUDENTS BASED ON:

IF YES, SPECIFY:

	Sex		Race	
Academic Interaction	Yes ○	No ○	Yes ○	No ○
Discipline	Yes ○	No ○	Yes ○	No ○
Positive Affect	Yes ○	No ○	Yes ○	No ○

POST-OBSERVATION CONFERENCE RECORD

SUMMARY FORM

Candidate: ____________________ S.S. #: ____________ Date: _______

Evaluator/#: ________________/_____ Times: _________ Visit: 1 2 3

Relative Strengths Exhibited:

Areas of Relative Need:

Recommendations:

Teacher Comments:

Evaluator Comments:

The signatures below **do not** necessarily indicate that the candidate agrees with the comments recorded by the observer. They simply verify that the conference has taken place.

______________________________	________________
Teacher's Signature	Date
______________________________	________________
Evaluator's Signature	Date

ELEMENTARY STUDENT QUESTIONNAIRE

GRADES 3-6

The Elementary Student Questionnaire is for use with grades three through six. Students anonymously answer thirty-four questions using a rating scale with the following response options: 1) almost never, 2) sometimes, 3) about half the time, 4) usually, 5) almost always and 6) don't know. Listed below are the questions and to the right of the question are a numeral and letter. The numeral and letter match the competency and indicator. This questionnaire is read aloud to the students by the evaluator.

1. Does your teacher explain to you what you are supposed to learn? IIA
2. Does your teacher show you how your school work is supposed to be learned? IIB
3. Can you ask your teacher questions about things you don't understand? IIB
4. When your teacher gives you work to do, does the teacher tell you when it is supposed to be finished? IVC
5. Does your teacher tell you why the things you study are important? IID
6. Does your teacher let you say things about what you are studying? IIC
7. Does your teacher notice when you are not paying attention in class? IID
8. Does your teacher teach this class by doing different things, such as group work, films, speakers, games and field trips? IIC
9. Does your teacher ask you to tell how you get answers? IIC
10. Does your teacher get you to think about how some things are alike? IIC
11. Does your teacher get you to think how something are different? IIC
12. Does your teacher help you to do your classwork while you are doing it? IIIB

13. Does your teacher give you homework? IIC

14. Does your teacher give you work to do by yourself during class time such as tests, worksheets? IIC

15. When you have questions about your work in this class, does the teacher help you to do the work correctly? IIIB

16. Does your teacher return your work to you quickly? IIIB

17. Does your teacher let your parents know how you are doing in school other than with report cards? IIIB

18. Is your teacher fair to all of the students in the class? IVA

19. Do you know how your teacher wants you to behave in the class? IVA

20. Does your teacher let you know when you are behaving correctly in this class? IVA

21. Does your teacher let you know when you are behaving incorrectly in this class? IVA

22. When your teacher is working with other students, does the teacher give you something to work on? IIC

23. Are things that you use in this class kept where they are easy to get when you need them? IVC

24. Does your teacher go (teach) too fast for you? IIB

25. Does your teacher go (teach) too slowly for you? IIB

26. Is your teacher on time for class? IVC

27. Does your teacher think you can learn? IVB

28. Does your teacher let you know when you should try harder? IVB

29. When you do well on your schoolwork, does your teacher tell you? IVB

30. Does your teacher want you to do the very best you can on your school work? IVB

31. Does your teacher have good things to say about students? IVB

32. Does your teacher talk badly about students in front of others? IVB

33. Have you learned a lot from this class? IIIC

34. Do you like to learn in this class? IIIC

SECONDARY STUDENT QUESTIONNAIRE

GRADES 7-12

The Secondary Student Questionnaire is for use with grades seven through twelve. Students answer the questions anonymously using a rating scale with the following response options: 1) almost never, 2) sometimes, 3) about half the time, 4) usually, 5) almost always, and 6) don't know. Listed below are the questions. The numeral and letter to the right of the question match the competency and indicator.

1. This teacher explains to us what we are supposed to learn. IIA
2. This teacher gives us clear directions for doing our school work. IIA
3. If I do not understand something that this teacher tells me, I feel that I can ask that it be explained again. IIB
4. This teacher allows me to make comments about what we are studying. IIC
5. When this teacher makes an assignment, the teacher tells us when it is supposed to be finished. IVC
6. This teacher requires students to complete their school work on time. IVC
7. This teacher tells us why the things we study are important. IID
8. This teacher notices when I am not paying attention in class. IID
9. This teacher uses different methods to teach, such as films, speakers, experiments, group/individual activities, games, and learning packets. IIC
10. This teacher asks us to explain how we get answers. IIC
11. This teacher comes to where I am working to check my classwork while I am doing it. IIIB
12. This teacher gives us work to do outside of class, such as projects, independent reading, or practice activities. IIC
13. This teacher gives us work to do by ourselves in class for practice. IIC

14.	This teacher gets us to think about the way things are similar.	IIC
15.	This teacher gets us to think about the way things are different.	IIC
16.	When I have questions about my work in this class, this teacher helps me to do the work correctly.	IIIB
17.	This teacher tells us how student work is to be graded.	IIIB
18.	This teacher returns our work to us quickly.	IIIB
19.	This teacher helps me learn how I can do a better job on my school work.	IIIB
20.	This teacher wants me to do the very best I can on my school work.	IVB
21.	This teacher is fair to all students.	IVA
22.	This teacher lets my parents know how I am doing.	IIIB
23.	This teacher keeps me informed about my progress.	IIIB
24.	This teacher wastes time in class.	IVC
25.	This teacher tells students how they should behave in this class.	IVA
26.	This teacher lets me know when I am behaving **correctly** in this class.	IVA
27.	This teacher lets me know when I am behaving **incorrectly** in this class.	IVA
28.	I have school work to do when the teacher is working with other students.	IIC
29.	The materials we use in this class are kept where they are easy to get when we need them.	IVC
30.	This teacher goes (teaches) too fast for me.	IIB
31.	This teacher goes (teachers) too slowly for me.	IIB
32.	This teacher is on time for class.	IVC
33.	This teacher thinks that I can learn.	IVB
34.	This teacher lets me know when I should try harder.	IVB

35. When I do well on my school work for this class, this teacher lets me know. IVB

36. This teacher has good things to say about students. IVB

37. This teacher talks badly about students in front of other people. IVB

38. I have learned about the subject being taught in this class. IIIC

39. I have enjoyed learning about the subject being taught in this class. IIIC

FOR VOCATIONAL EDUCATION STUDENTS ONLY

40. This teacher explains how classroom activities relate to business and industry. IB

41. This teacher lets me know what the safety rules are for this class. IVC

42. This teacher makes sure that all safety rules are followed. IVC

PRINCIPAL QUESTIONNAIRE

The Principal Questionnaire is completed for each teacher by the person who is responsible for the teacher's personnel evaluation. In most instances this is the school principal, but it may be an assistant principal or a supervisor.

The questionnaire has seventeen questions for the regular classroom teacher, eighteen questions for the vocational education teacher, and nineteen questions for the special populations teacher who is required by law to do Individual Educational Programs (IEPs). Each question is an indicator under a competency in the evaluation system.

Principals are asked to respond using a ten-point scale as shown below:

Career Ladder Status	Rating Score	200-800
III	3+	800
	3	750
	3-	725
II	2+	699
	2	650
	2-	625
I	1+	599
	1	450
	1-	300
Below Level I	0	200

TENNESSEE CAREER LADDER
TEACHER SYSTEM
PRINCIPAL QUESTIONNAIRE

CANDIDATE NAME:____________________________ SOC. SEC #:____________________

SCHOOL NAME:____

DIRECTIONS: For each question, place an "X" in the box which best describes your assessment of the teacher for placement on the Career Ladder. Rating a teacher should not be dependent on the level for which the teacher is applying. For example, a teacher applying for Career Level II may receive a Career Level III score. The rating should be made on the teacher's performance, not on the application level. At the end of each question, you will find reference to the domain and indicator which are measured by that question.

	Level 3			Level 2			Level 1			Below	Don't Know
	3+	3	3-	2+	2	2-	1+	1	1-	0	
1. Establishes appropriate instructional goals and related objectives consistent with the curriculum (IA)											
2. Prepares instructional plans and materials incorporating principles of effective instruction (IB)											
3. Creates, selects or modifies instructional plans and materials to accommodate learner instructional levels (IC)											
4. Provides a clear description of the learning task and its content (IIA)											
5. Monitors learner understanding and reteaches as necessary (IIB)											
6. Provides learners appropriate practice and review (IIC)											
7. Establishes and maintains learner involvement in the learning task (IID)											
8. Uses information about learner performance to improve the instructional process (IIIA)											

	Level 3			Level 2			Level 1			Below	Don't Know
	3+	3	3-	2+	2	2-	1+	1	1-	0	
9. Reports learner status and progress to learners and their parents (IIIB)											
10. Establishes and maintains appropriate learner behavior (IVA)											
11. Establishes and maintains a classroom climate conducive to learning (IVB)											
12. Makes effective use of classroom resources (IVC)											
13. Improves professional skills and knowledge (VA)											
14. Takes a leadership role in improving education (VB)											
15. Communicates oral information effectively (IIA)											
16. Improves learners' academic achievement in relevant subject areas (IIIC)											
17. Improves learners' attitudes toward learning(IIIC)											
18. Serves as a liaison for home, school and community **(Vocational Education)** (VD)											
19. Integrates and facilitates individual education programs (IA) **(Special Education only)**											
20. Communicates effectively with professionals,paraprofessionals, and parents **(Special Education Only)**(V)											

TEACHER/EVALUATOR DIALOGUE

Planning Focus

1. How do you decide what to teach? Briefly outline your instructional program for the year. Discuss your goals and objectives for your unit plan and lesson plan.

NOTES:

SCORES:

	unsatisfactory	below average	average	outstanding	distinguished
1. Use of goals and objectives related to curriculum	1	2	3	4	5
2. Use of scope and sequence	1	2	3	4	5
3. Addressing of basic content skills	1	2	3	4	5
4. Addressing higher level thinking skills	1	2	3	4	5

TEACHER/EVALUATOR DIALOGUE

Planning Focus

2. How do you plan for your teaching strategies and the use of your instructional materials?

NOTES:

SCORES:

	unsatisfactory	below average	average	outstanding	distinguished
5. Logic and completeness of instructional plans	1	2	3	4	5
6. Use of introductory/ motivational procedures	1	2	3	4	5
7. Use of material/media related to goals/objectives	1	2	3	4	5
8. Use of material/media appropriate for students	1	2	3	4	5

TEACHER/EVALUATOR DIALOGUE

Planning Focus

3. How do you adjust for student differences in your planning?

NOTES:

SCORES:	unsatisfactory	below average	average	outstanding	distinguished
9. Use of varying instructional plans to adjust for student differences	1	2	3	4	5
10. Use of varying instructional grouping to adjust for student differences	1	2	3	4	5
11. Use of varying materials/ activities to adjust for student differences	1	2	3	4	5

TEACHER/EVALUATOR DIALOGUE

Teaching Strategies Focus

1. How do you prepare students for what they will be doing in the classroom and how they are to do it. How do you present subject matter to students?

NOTES:

SCORES:	unsatisfactory	below average	average	outstanding	distinguished
1. Demonstrates knowledge of correct/current information	1	2	3	4	5
2. Pacing of presentation according to difficulty of material	1	2	3	4	5
3. Relationship and integration of task or content	1	2	3	4	5
4. Use of clear directions, explanations and examples	1	2	3	4	5
5. Use of appropriate and correct language model	1	2	3	4	5

TEACHER/EVALUATOR DIALOGUE

Teaching Strategies Focus

2. How do you keep track of student learning during a lesson? What do you do if students do not understand a lesson?

NOTES:

SCORES:

	unsatisfactory	below average	average	outstanding	distinguished
6. Demonstration of knowledge and use of student instructional levels	1	2	3	4	5
7. Monitoring and checking for student understanding	1	2	3	4	5
8. Adjustment for student differences	1	2	3	4	5
9. Attending to students' needs	1	2	3	4	5

TEACHER/EVALUATOR DIALOGUE

Teaching Strategies Focus

3. After you have introduced and taught a lesson, how do you provide for student practice and review?

NOTES:

SCORES:	unsatisfactory	below average	average	outstanding	distinguished
10. Provision for supervised and independent practice	1	2	3	4	5
11. Required use of skills/ concepts	1	2	3	4	5
12. Provision for practice in higher order thinking skills	1	2	3	4	5
13. Provision for appropriate review	1	2	3	4	5

TEACHER/EVALUATOR DIALOGUE

Teaching Strategies Focus

4. How do you keep students involved and participating in learning activities?

NOTES:

SCORES:

	unsatisfactory	below average	average	outstanding	distinguished
14. Identification of importance of learning to students	1	2	3	4	5
15. Use of varied and effective strategies for student involvement	1	2	3	4	5
16. Provision for active student participation	1	2	3	4	5
17. Adaptation/adjustment of lesson for student involvement	1	2	3	4	5

TEACHER/EVALUATOR DIALOGUE

Evaluation Focus

1. How do you evaluate your students' progress? How do you evaluate what you teach?

NOTES:

SCORES:

		unsatisfactory	below average	average	outstanding	distinguished
1.	Measuring achievement of goals and objectives	1	2	3	4	5
2.	Monitoring student progress	1	2	3	4	5
3.	Use of assessment	1	2	3	4	5
4.	Use of data to assess curriculum and instruction	1	2	3	4	5

TEACHER/EVALUATOR DIALOGUE

Evaluation Focus

2. When do you evaluate student progress and what do you do with your evaluation results?

NOTES:

SCORES:

		unsatisfactory	below average	average	outstanding	distinguished
5.	Accuracy and timeliness of assessment records	1	2	3	4	5
6.	Use of performance data to adjust instruction	1	2	3	4	5
7.	Reporting of student progress to students, parents, and appropriate others	1	2	3	4	5
8.	Use/understanding evaluation data	1	2	3	4	5

TEACHER/EVALUATOR DIALOGUE

Evaluation Focus

3. What do you expect of students in your class(es) and how do you communicate that to them? How well do students perform in your class? How do you know what progress they have made in knowledge, skills, and attitudes?

NOTES:

SCORES:

		unsatisfactory	below average	average	outstanding	distinguished
9.	Evidence of high expectations for student learning	1	2	3	4	5
10.	Provision of appropriate grouping for student learning	1	2	3	4	5
11.	Evidence of student academic achievement commensurate with background/ability of students	1	2	3	4	5
12.	Evidence of improvement in student attitude toward learning	1	2	3	4	5

PROFESSIONAL DEVELOPMENT AND LEADERSHIP SUMMARY

INTRODUCTION

Professional Leadership is one of the Domains of Competence which support both the focus and foundation of the Career Ladder evaluation system. Items from the Principal's Questionnaire and the Professional Development and Leadership Summary (PDL) comprise the data sources for this domain. The PDL is a written summary of individual activities in 10 or 11 categories submitted by the candidate under two major areas: Professional Development and Leadership. This part of the Instrument Section provides information and instructions necessary to report the PDL activities.

Summative descriptions of the professional development and leadership activities are required of all upper level Career Ladder candidates. If a candidate does not submit the PDL Summary by the close of the first state conducted visit, the evaluation process is terminated. The PDL Summary becomes part of the candidate's evaluation file and is not returned to the candidate. To assist in understanding and in completing the PDL Summary, the following sections are included: **RATING SCALE; GENERAL INSTRUCTIONS; CATEGORIES OF ACTIVITIES;** and **DIRECTIONS FOR COMPLETING FORMS.**

RATING SCALE

Each summative description of an activity is rated on a five point quality scale: 1) unsatisfactory, 2) below average, 3) average, 4) outstanding, and 5) distinguished. Ratings are made according to the degree of correlation to the category description, relationship to the educational environment, and application to the teaching setting. Further, PDL descriptions are rated based on statements of the rationale, purpose/goal, content, and scope of the activity, including the relative time and effort spent on the activity. Ratings are also based on the candidate's evaluation of the activity, specific benefits, outcomes, and follow up.

GENERAL INSTRUCTIONS

General instructions which apply when writing the PDL Summary are given below:

- The PDL is the only required written data source and must be completed before the beginning of the first state conducted visit. Submission of the PDL is due on the day of the first state conducted visit.

- Activity descriptions should be exact and detailed. The clarity and specificity of the information are directly related to the rating.

- The beginning date of an activity may occur during any year of employment. However, activities must have been completed during the last five teaching years. If the activity is ongoing, the opportunity to implement and evaluate the activity should have occurred within the last five years.

  Exceptions:

  The advanced degree may be used regardless of the date obtained.

  A break in service extends the five year requirement beyond the past five calendar years.

- If the information on beginning and ending dates is not recorded on the PDL forms, the activity will not be rated.

- Summaries of each activity must follow the provided PDL format. However, the forms may be reproduced or retyped to allow additional space for answers.

- Answers may be completed on a separate sheet and must be labeled.

- The candidate's signature and date indicate verification for engagement in the activity as well as confirm the accuracy of all written information.

- Submitted forms without signature and date will not be rated.

- Submitted forms should be arranged with Professional Development activities placed first followed by Leadership activities.

- The candidate is not required to staple the PDL activity forms.

- No documentation should be included. Documentation (such as letters, transcripts, examples, and products) will be returned to the candidate by the evaluator collecting the PDL Summary or destroyed upon arrival at the Career Ladder office.

- Only one activity for each category under Professional Development and one activity for each category under Leadership may be submitted. Additional activities will be returned to the candidate by the evaluator or destroyed when received in the Career Ladder office.

- For **General Education**: A completed PDL includes one written activity for each of the five Professional Development categories and one for each of the five Leadership categories for a total of 10 activities.

- For **Special Populations I**: A completed PDL includes one written activity for each of the five Professional Development categories and one for each of the six Leadership categories for a total of 11 activities.

  Special Populations I includes teachers of students in the following programs:

  Group 1:

  Chapter I
  Vocational Improvement Program
  English as a Second Language

  Group 2:

  Vocational Advancement Program
  Intellectually Gifted
  Resource, Mildly Handicapped
  Visually Impaired
  Hearing Impaired

- For **Vocational**: A completed PDL includes one written activity for each of the five Professional Development categories and one for each of the six Leadership categories for a total of 11 activities.

  Vocational includes teachers of students in the following programs:

  Trade and Industry
  Industrial Arts
  Vocational Office
  Agriculture
  Marketing-Distributive
  Health Occupations
  Occupational Home Economics
  Consumer Home Economics

Exceptions:

- **Vocational** teachers of students in the following programs complete the same PDL requirements as **General Education** teachers, (that is, one for each of the five Professional Development categories and one for each of the five Leadership categories for a total of 10 activities):

  Business Education
  General Agriculture
  General Home Economics
  Career Education
  Prevocational

CATEGORIES OF ACTIVITIES

PROFESSIONAL DEVELOPMENT

Professional Development comprises a variety of activities undertaken by a candidate to increase professional skills, attitudes, and knowledge. Activities also relate to the candidate's educational assignment and/or enhance advancement within the profession. As a result, improved instruction or quality of services to children is expected.

The categories for Professional Development are as follows:

PD #1 Obtains Graduate Degree(s) and/or Takes Courses

Graduate degrees or coursework utilized for this category must relate to your professional assignment and/or professional advancement. In order to complete the requirements for this category, choose one of the following for the basis of your answer:

- Degree above bachelor's degree (no time limit). List the major area of study and date received.

or

- Summary of professional coursework engaged in beyond graduate degree and taken during the last five teaching years.

or

- Summary of professional coursework engaged in during the last five teaching years. This coursework may or may not have been in pursuit of a degree.

or

- Single professional course taken during the last five teaching years.

PD #2 Participates in Professional Development Activities

Describe one professional development activity in which you participated for the purpose of enhancing your instructional performance. This category excludes required inservice activities and training in the Tennessee Instructional Model (TIM) if used for fast track purposes during 1984-85. Staff development activities designed to improve (increase) knowledge, skills, and attitudes are appropriate. Examples of acceptable activities include continuing education courses, educational (professional) conventions and conferences, workshops, and TIM (if not used for fast track).

PD #3 Uses Ideas from Professional Books, Journals, and Professional Organizations to Enhance Classroom Instruction

Describe one idea taken from a professional resource or resources and identify the derivative source. The source should not be in use as a textbook in your school. Include an explanation of how the idea was adapted to your teaching assignment.

PD #4 Tries New Methods/Approaches in the Classroom and Evaluates Success

Describe one new method/approach implemented in your classroom in order to improve your teaching performance. Explain the evaluation method and results. Give the source of the method/approach. Examples of sources include your own idea, idea(s) from a peer or professional magazine, or idea(s) received from a professional development activity.

PD #5 Seeks and Utilizes Community Resources to Enhance Classroom Instruction

Describe one activity which improved classroom instruction by seeking and utilizing community resources. Community resources may be people, places, and materials and may reach beyond the local area. Explain the relationship to the curriculum, your instructional goals and objectives, and student needs.

LEADERSHIP

Leadership includes a variety of activities undertaken by a candidate for the purpose of contributing to the professional growth of others. Activities also relate to improving education for students within the school/school system environment.

The categories for Leadership are as follows:

L #1 Conducts Workshops or Training Sessions for Peers

Describe a professional workshop or training session which you have conducted for peers in the school, school system, district, or state. Explain how the participants used the knowledge, skills, or attitudes presented by you. Training of student teachers may be used in this category.

L #2 Creates Materials or Programs and Shares with Peers

Describe one example of materials or one program which you originated and shared with peers. This activity may refer to committee, team work, or individual effort. Explain how the recipients used the materials or program, and how you assisted them. Describe the extent of the resultant sharing/use by others.

L #3 Holds a Leadership Position in the School/School System or Educational Organization

Describe one voluntary or elected leadership position for which you did not receive payment. This category covers grade or group chairperson positions which address peer groups, educational honors, or a leadership position held in an organization or committee whose primary function is educational.

L #4 Promotes Parent/Community Interest in the School

Describe one activity in which you encouraged and improved parent/community interest and involvement in the school. This category includes sponsorship of school club activities. Explain how this activity promoted interest in the school and describe the extent of resultant parent/community involvement.

L #5 Initiates Activities/Projects in the School

Describe one activity/project which you were responsible for starting in your school. This category may or may not refer to an activity which you designed or created. This activity should involve peers but may include student groups if impact is school wide. Specify the extent of peer and/or student involvement.

L #6 Communicates Effectively with Professionals, Paraprofessionals, and/or Parents (Special Populations I & II Only)

Describe one activity which refers to a method or technique utilized to enhance communication with professionals and/or parents. Activities may revolve around the multi-disciplinary team (M-team) process, conferences, and/or procedures used for the purpose of planning for and meeting student needs. If your activity involves communication with parents, integration of home and school objectives should be discussed.

L #7 Establishes and Uses Advisory Groups (Vocational Only)

Describe one activity of your program craft or advisory committee. The school or system advisory committee should not be used for this category. Explain how you were involved in the activity and the resulting benefits to students, school, employers, or self.

DIRECTIONS FOR COMPLETING PROFESSIONAL DEVELOPMENT FORM

I. Print your **name** and **social security number** as it appears on your application.

II. Circle one **category**. Use one form per category.

III. Write a **descriptive title** for your activity. Examples are Math Project to Teach Graphing, Post Graduate Studies, Workshop on Assertive Discipline.

IV. Indicate **beginning** and **ending dates**. Dates should be written as:

day/month/year or **month/year** or **season/year**

The ending date must have occurred during the last five teaching years unless: (1) there has been a break in service; or (2) the activity is ongoing. If the activity is ongoing, an explanation must be given.

V. Write **approximate total time** in hours devoted to this activity. Include the total number of hours spent in planning, implementing, and evaluating. Write a statement about when you worked on this activity. Examples include the following: after school, weekends, summers, and release time. For PD #1, degrees and coursework may be listed in semester or quarter hours.

VI. State your **rationale** or reason for undertaking this activity. Indicate relationship to curriculum, teaching assignment, and/or educational area. Identify specific need(s). State your **professional goal/purpose**.

VII. Explain your **role**. To describe your role as student, initiator, or participant is insufficient. Explain your **responsibilities**.

VIII. **Describe the details of the activity.** Include the scope, implementation strategies, and adaptations to your teaching assignment. Write the evaluation method for the activity whether formal or informal, including the results of evaluation, outcomes, and follow up. Evaluation method should be included for categories PD3, PD4, and PD5.

IX. Explain specific **benefits to you** in your professional role. Include how the knowledge attained was used and what changes/improvements occurred, as well as specific outcomes.

X. Explain specific **benefits to your students** or **school system**. Include outcomes, conclusions, and resultant changes/improvements.

Note: Forms without signature and date will not be rated.

PROFESSIONAL DEVELOPMENT FORM

I. NAME:______________________________ S.S#:____________________

II. CATEGORY: 1 2 3 4 5

III. ACTIVITY TITLE:

IV. BEGINNING DATE:______________ ENDING DATE:________________
IF ONGOING, EXPLAIN:

V. APPROXIMATE TIME:________ HOURS
STATEMENT ABOUT WHEN YOU WORKED ON THIS ACTIVITY:

VI. RATIONALE AND PROFESSIONAL GOAL

VII. ROLE AND RESPONSIBILITIES

VIII. ACTIVITY DESCRIPTION

IX. BENEFITS TO YOU IN INSTRUCTIONAL ROLE

X. BENEFITS TO YOUR STUDENTS, OR CHANGES/IMPROVEMENTS MADE TO YOUR SCHOOL/SCHOOL SYSTEM

I verify that I personally engaged in this activity and that all the information on this form is accurate.

______________________________	________________
Teacher's Signature	**Date**

DIRECTIONS FOR COMPLETING LEADERSHIP FORM

I. Print your **name** and **social security number** as it appears on your application.

II. Circle one **category**. Use one form per category.

III. Write a **descriptive title** for your activity. Examples are Assertive Discipline Training, Chairman-Textbook Committee, Developer of Science Computer Software Packet.

IV. Indicate **beginning** and **ending dates**. Dates should be written as:

day/month/year or **month/year** or **season/year**

The ending date must have occurred during the last five teaching years unless: (1) there has been a break in service; or (2) the activity is ongoing. If the activity is ongoing, an explanation must be given.

V. Write **approximate total time** in hours devoted to this activity. Include the total number of hours spent in planning, implementing, attending meetings, evaluating and making recommendations. Write a statement about when you worked on this activity. Examples include the following: after school, weekends, summers, and release time.

VI. State your **rationale** or reason for undertaking this activity. Indicate relationship to curriculum, teaching assignment, and/or educational area. Identify specific need(s).

State your **professional goal/purpose**.

VII. Explain your **role**. To describe your role as presenter, coordinator, or representative is insufficient.

Explain your **responsibilities**.

VIII. **Describe the details of the activity.** Include content, frequency of occurrence, method of evaluation for activity, whether formal or informal, including the results of evaluation, outcomes, and follow up. Evaluation method should be included in the descriptions for all Leadership categories.

IX. Explain specific **benefits to you/your students or the school/school system** as a result of your participation in this activity.

Note: Forms without signature and date will not be rated.

LEADERSHIP FORM

I. NAME:______________________________ S.S#:____________________

II. CATEGORY: 1 2 3 4 5 6 7

III. ACTIVITY TITLE:

IV. BEGINNING DATE:________________ ENDING DATE:__________________
IF ONGOING, EXPLAIN:

V. APPROXIMATE TIME:________ HOURS
STATEMENT ABOUT WHEN YOU WORKED ON THIS ACTIVITY:

VI. RATIONALE AND PROFESSIONAL GOAL

VII. ROLE AND RESPONSIBILITIES

VIII. ACTIVITY DESCRIPTION

IX. BENEFITS TO YOU/YOUR STUDENTS OR SCHOOL/SCHOOL SYSTEM

I verify that I personally engaged in this activity and that all the information on this form is accurate.

______________________________ ____________________

Teacher's Signature Date

Appendix G: How Information Is Obtained About Competencies and Indicators

GENERAL EDUCATION
HOW INFORMATION IS OBTAINED
ABOUT COMPETENCIES AND INDICATORS

	CLASSROOM OBSERVATION	DIALOGUE	PRINCIPAL QUESTIONNAIRE	STUDENT QUESTIONNAIRE	PROFESSIONAL DEVELOPMENT AND LEADERSHIP SUMMARY	WRITTEN TEST	CONSENSUS
I. Prepares for instruction effectively.						X	X
A. Establishes appropriate instructional goals and related objectives consistent with the curriculum.	X	X	X				
B. Prepares instructional plans and materials incorporating principles of effective instruction.	X	X	X				
C. Creates, selects or modifies instructional plans and materials to accommodate learner instructional levels.	X	X	X				
II. Uses teaching strategies and procedures appropriate to the content, objectives and learners.						X	X
A. Provides a clear description of the learning task and its content.	X	X	X	X			
B. Monitors learner understanding and reteaches as necessary.	X	X	X	X			
C. Provides learners appropriate practice and review.	X	X	X	X			
D. Establishes and maintains learner involvement in the learning task.	X	X	X	X			
III. Uses evaluation to improve instruction.						X	X
A. Uses information about learner performance to improve the instructional process.		X	X				
B. Reports learner status and progress to learners and their parents.	X	X	X	X			
C. Improves learner performance.		X	X	X			
IV. Manages classroom activities effectively.						X	X
A. Establishes and maintains appropriate learner behavior.	X		X	X			
B. Establishes and maintains a classroom climate conducive to learning.	X	X	X	X			
C. Makes effective use of classroom resources (e.g., personnel, time, materials, facilities).	X		X	X			
V. Establishes and maintains a professional leadership role.							
A. Improves professional skills and knowledge.			X		X		
B. Takes a leadership role in improving education.			X		X		
C. (Screening only) Performs professional responsibilities efficiently.			X				
VI. (Screening only) Communicates effectively.						X	
A. Writes clearly and correctly.			X			X	
B. Reads professionally relevant literature/materials with comprehension.			X			X	

SPECIAL EDUCATION

HOW INFORMATION IS OBTAINED

ABOUT COMPETENCIES AND INDICATORS

	CLASSROOM OBSERVATION	DIALOGUE	PRINCIPAL QUESTIONNAIRE	STUDENT QUESTIONNAIRE	PROFESSIONAL DEVELOPMENT AND LEADERSHIP SUMMARY	WRITTEN TEST	CONSENSUS
I. Prepares for instruction effectively.						X	X
A. Establishes appropriate instructional goals and related objectives consistent with the curriculum.	X	X	X				
B. Prepares instructional plans and materials incorporating principles of effective instruction.	X	X	X				
C. Creates, selects or modifies instructional plans and materials to accommodate learner instructional levels.	X	X	X				
II. Uses teaching strategies and procedures appropriate to the content, objectives and learners.						X	X
A. Provides a clear description of the learning task and its content.	X	X	X	X			
B. Monitors learner understanding and reteaches as necessary.	X	X	X	X			
C. Provides learners appropriate practice and review.	X	X	X	X			
D. Establishes and maintains learner involvement in the learning task.	X	X	X	X			
III. Uses evaluation to improve instruction.						X	X
A. Uses information about learner performance to improve the instructional process.		X	X				
B. Reports learner status and progress to learners and their parents.	X	X	X	X			
C. Improves learner performance.		X	X	X			
IV. Manages classroom activities effectively.						X	X
A. Establishes and maintains appropriate learner behavior.	X		X	X			
B. Establishes and maintains a classroom climate conducive to learning.	X	X	X	X			
C. Makes effective use of classroom resources (e.g., personnel, time, materials, facilities).	X		X	X			
V. Establishes and maintains a professional leadership role.							
A. Improves professional skills and knowledge.			X		X		
B. Takes a leadership role in improving education.			X		X		
C. (Screening only) Performs professional responsibilities efficiently.			X				
D. Communicates effectively with professionals, paraprofessionals and parents.			X		X		
VI. (Screening only) Communicates effectively.						X	
A. Writes clearly and correctly.			X			X	
B. Reads professionally relevant literature/materials with comprehension.			X			X	

VOCATIONAL EDUCATION

HOW INFORMATION IS OBTAINED

ABOUT COMPETENCIES AND INDICATORS

	CLASSROOM OBSERVATION	DIALOGUE	PRINCIPAL QUESTIONNAIRE	STUDENT QUESTIONNAIRE	PROFESSIONAL DEVELOPMENT AND LEADERSHIP SUMMARY	WRITTEN TEST	CONSENSUS
I. Prepares for instruction effectively.						X	X
A. Establishes appropriate instructional goals and related objectives consistent with the curriculum.	X	X	X				
B. Prepares instructional plans and materials incorporating principles of effective instruction.	X	X	X	X			
C. Creates, selects or modifies instructional plans and materials to accommodate learner instructional levels.	X	X	X				
II. Uses teaching strategies and procedures appropriate to the content, objectives and learners.						X	X
A. Provides a clear description of the learning task and its content.	X	X	X	X			
B. Monitors learner understanding and reteaches as necessary.	X	X	X	X			
C. Provides learners appropriate practice and review.	X	X	X	X			
D. Establishes and maintains learner involvement in the learning task.	X	X	X	X			
III. Uses evaluation to improve instruction.						X	X
A. Uses information about learner performance to improve the instructional process.		X	X				
B. Reports learner status and progress to learners and their parents.	X	X	X	X			
C. Improves learner performance.		X	X	X			
IV. Manages classroom activities effectively.						X	X
A. Establishes and maintains appropriate learner behavior.	X		X	X			
B. Establishes and maintains a classroom climate conducive to learning.	X	X	X	X			
C. Makes effective use of classroom resources (e.g., personnel, time, materials, facilities).	X	X	X	X			
V. Establishes and maintains a professional leadership role.							
A. Improves professional skills and knowledge.			X		X		
B. Takes a leadership role in improving education.			X		X		
C. (Screening only) Performs professional responsibilities efficiently.			X				
D. Serves as a liaison for home, school and community.			X	X	X		
VI. (Screening only) Communicates effectively.						X	
A. Writes clearly and correctly.			X			X	
B. Reads professionally relevant literature/materials with comprehension.			X			X	